

1-1998

## Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science

Michael J. Saks

Follow this and additional works at: [https://repository.uchastings.edu/hastings\\_law\\_journal](https://repository.uchastings.edu/hastings_law_journal)



Part of the [Law Commons](#)

---

### Recommended Citation

Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069 (1998).

Available at: [https://repository.uchastings.edu/hastings\\_law\\_journal/vol49/iss4/9](https://repository.uchastings.edu/hastings_law_journal/vol49/iss4/9)

This Article is brought to you for free and open access by the Law Journals at UC Hastings Scholarship Repository. It has been accepted for inclusion in Hastings Law Journal by an authorized editor of UC Hastings Scholarship Repository.

# Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science

by  
MICHAEL J. SAKS\*

In scientific inquiry it becomes a matter of duty to expose a supposed law to every possible kind of verification, and to take care, moreover, that this is done intentionally, and not left to a mere accident.<sup>1</sup>

## Introduction

Scientific expert opinion is a specie of evidence that presents the courts with a challenging set of paradoxes and dilemmas. Consider: Scientific expert testimony is usually welcomed because it promises to help resolve trial uncertainties with knowledge that is otherwise unavailable to lay jurors and judges.<sup>2</sup> Yet it is those same lay judges who must serve as evidentiary gatekeepers,<sup>3</sup> admitting only valid sci-

---

\* Edward F. Howrey Professor, University of Iowa College of Law; Ph.D., Ohio State University, M.S.L., Yale Law School. This article grew out of work done in the course of preparing my share of DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS, & JOSEPH SANDERS, *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* (1997). I wish to thank: John "Max" Wilkinson and Alec Hillbo for their excellent research assistance; colleagues David Faigman and David Kaye for their helpful comments and criticisms of portions of this work; and Michael Risinger, who should be regarded as a co-author of certain sections of the article. See *infra* notes 18-21 and 117-153, and accompanying text. Portions of this paper were presented at the Annual Tom Krauss Memorial Bitemark Breakfast Address of the American Society of Forensic Odontology, New York City, February 21, 1997, and at a faculty seminar at the Georgetown University Law Center, November 12, 1996.

1. T. H. HUXLEY, *The Method of Scientific Investigation*, in *SCIENCE: METHOD AND MEANING* (Samuel Rapport & Helen Wright eds., 1974).

2. The Federal Rules of Evidence have dispensed with a "beyond the ken" requirement as a *pre-requisite* for admissibility, in favor of a "helpfulness" standard. See FED. R. EVID. 702. But the fact that non-expert judges and jurors are not conversant with a subject remains the principal reason expert testimony is offered. See FED. R. EVID. 702 advisory committee's note.

3. That judges have a duty to screen scientific evidence is implicit in any test for admissibility, see FED. R. EVID. 104, and is made explicit in the references to "gatekeep-

entific evidence while excluding "junk science." How can they distinguish one from the other? How can evidence, at one and the same time, be beyond judges' knowledge and yet amenable to their intelligent appraisal? The problem is not helped by the cultural divide that places judges and lawyers among society's most improbable candidates to be assigned the task of assessing claims of purported science.<sup>4</sup> How can judges be expected to know what questions need to be asked and what to make of the answers? This paradox is one of the roots of at least a century of debate over the process of receiving proffered scientific evidence.<sup>5</sup> It has led the courts to experiment with a variety of tests and procedures.<sup>6</sup> So far, however, the law does not seem to have invented a notably successful resolution to the problem.<sup>7</sup>

At first blush, no evidence would seem better suited than scientific evidence to focus courts' attention on the truth-finding function

ing" in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 589 n.7, 597 (1993).

4. For a description of the cultural divisions between law and science, see ANDRE MOENSSSENS ET AL., *SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES* 10 (4th ed. 1995) ("[A]s a group, lawyers display an appalling degree of scientific illiteracy . . ."); Peter H. Schuck, *Multi-Culturalism Redux: Science, Law and Politics*, 11 *YALE L. & POL'Y REV.* 1 (1993); C. P. SNOW, *THE TWO CULTURES AND THE SCIENTIFIC REVOLUTION* (1959) (discussing the wide gulf between the two cultures of science and the humanities).

5. See, e.g., Lee M. Friedman, *Expert Testimony, Its Abuse and Reformation*, 19 *YALE L.J.* 247, 247 (1910):

The position of an expert on the witness stand, who does not testify either to what he has observed or knows as fact but expresses merely his opinion as to a situation or on facts which have been established by other witnesses, is anomalous in Anglo-Saxon law. It was to be expected that former generations of judges and lawyers trained in older precedents and practices who recognized the appearance in the courts of an expert witness as an innovation would look with suspicion and doubt on such testimony. While the principles on which such evidence is introduced have come to be well recognized and while the profession no longer has any reservation in approving theoretically of the use of expert testimony, yet, on the other hand, there is a constant complaining and mistrust on the part of judges, juries and lawyers of the expert witness.

For an example of the view of expert testimony at the turn of the century, see, for example, Clemens Herschel, *Services of Experts in the Conduct of Judicial Inquiries*, 21 *AM. L. REV.* 571, 571-72 (1887):

Many judges have expressed their thorough dissatisfaction with the prevalent method . . . of making use of the services of experts in the conduct of judicial inquiries. . . . No judge has, in recent times, said aught in praise of the system, so far as the writer has been able to observe. Law writers are equally condemnatory of the system, and severe in their reflections upon the product of that system—the expert.

6. See *infra* part II.

7. See, e.g., *Developments in the Law—Confronting the New Challenges of Scientific Evidence*, 108 *HARV. L. REV.* 1481, 1509-16, 1583-1604 (1995) (summarizing various proposed and adopted solutions to the problems posed by scientific evidence).

of trials. Yet, at least in regard to the kinds of assertedly scientific evidence discussed in this article, the courts seem to have treated the expert evidence more like magic than science. In terms of judicial methodology, they might as well have been validating Cotton Mather's expertise for identifying witches. These courts appear to have been looking at the totemic properties of science (that is, the imprimatur that the institution of science could place on evidence) rather than anything that scientists rely on in assessing scientific claims (the data and logic and testing of empirical claims).

The present article examines the performance of the courts in evaluating one important and ubiquitous genre of expert evidence: forensic identification science.<sup>8</sup> How did those courts scrutinize the major forensic identification sciences in their precedent-setting encounters? Were these courts seeking primarily to enhance the truth-seeking capacity of trials? Or something else, such as improving the capacity of courts to quiet the jury's and the public's doubts? What are the lessons that this past offers us for the future?

For many areas of science, revisiting the courts' historical performance is virtually necessitated by the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,<sup>9</sup> which requires federal judges to stop merely taking the word of experts for the general acceptance of what they came to court to offer,<sup>10</sup> and to begin reviewing the evidence underlying that evidence for its validity. Under *Daubert*, no purportedly scientific statement can be valid unless it is capable of being empirically tested; no theory can be scientific unless it is subject to falsification. Under *Daubert*, past admission is not sufficient to assure the continued admissibility of any purported scientific evidence.<sup>11</sup> Continued admissibility—or, indeed, continued exclusion—depends instead on the quality of the evidence about the evidence and the quality of judicial reasoning about that evidence.

---

8. Forensic identification science is defined *infra* Part II.A. The category includes identification by fingerprints, handwriting, voice, footprints, DNA, fingernail clippings, toolmarks, bitemarks, and so on.

9. 509 U.S. 579 (1993).

10. The test under *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923), did not specify how a court was to find general acceptance. In practice, judges usually were content to rely on the testimony of the proffered expert about that issue. As we shall see, even when a high degree of controversy in the literature (not just in the courtroom) is brought to the court's attention—which would seem by definition to refute the claim of general acceptance—the few witnesses supporting admission sometimes were enough to establish “general acceptance.”

11. “Although the *Frye* decision itself focused exclusively on ‘novel’ scientific techniques, we do not read the requirements of Rule 702 to apply specially or exclusively to unconventional evidence. Of course, well-established propositions are less likely to be challenged than those that are novel, and they are more handily defended.” *Daubert*, 509 U.S. at 592 n.11.

*Daubert* invites so new a look at old scientific evidence, and from such a different perspective than most earlier tests, that many scientific evidence precedents are now vulnerable to reconsideration and reversal. Some, of course, are more vulnerable than others. Under *Daubert* review, forensic identification science turns out to be among the most vulnerable.<sup>12</sup> Indeed, under *Daubert* review, several federal courts already have held handwriting identification expertise to be non-science<sup>13</sup> and another has held hair identification expertise inadmissible.<sup>14</sup>

Part I of this article presents a brief discussion of the major legal tests for the admissibility of scientific evidence. While such ground has been trod repeatedly in discussions of law and science, here emphasis is given to explicating several myths that routinely accompany discussions of these tests, and which are relevant to this article's analysis of scientific evidence admissibility. In this fresh light, the courts' difficulties may become more clear. Part II is a brief intellectual history of the forensic identification sciences, focusing on the puzzle of the arrested development of scientific research in these fields. Part III summarizes the histories of the courts' pivotal encounters with several major forensic identification sciences over much of the 20th century and explicates how the courts analyzed each of them.<sup>15</sup> The paper concludes with a discussion of the lessons that might be learned from these encounters about the mutual impact of the law and purported science and for judicial regulation of the

---

12. Some commentators are skeptical about *Daubert's* likely impact. See, e.g., Randolph Jonakait, *The Meaning of Daubert and What it Means for Forensic Science*, 15 CARDOZO L. REV. 2103 (1994). Other illuminating discussions of the move from *Frye* to *Daubert*, particularly as it may affect forensic science, include: Paul C. Giannelli, *Forensic Science: Frye, Daubert and the Federal Rules*, 29 CRIM. L. BULL. 428 (1993); Paul C. Giannelli, *Daubert: Interpreting the Federal Rules of Evidence*, 15 CARDOZO L. REV. 1999 (1994); and "Focus" articles in 1 SHEPARD'S SCIENTIFIC AND EXPERT EVIDENCE QUARTERLY (Winter, 1994).

13. See *infra* notes 138-150 and accompanying text. This does not mean that these courts did not find ways to admit the testimony of forensic document examiners ("FDEs").

14. As noted by *Williamson v. Reynolds*, 904 F. Supp. 1529, 1558 (E.D. Okla. 1995) (quoting *Daubert*, 509 U.S. at 592 n. 11):

The state of the art of hair analysis has not reached a level of certainty to permit such testimony. Although the hair expert may have followed procedures accepted in the community of hair experts, the human hair comparison results in this case were, nonetheless, scientifically unreliable. This court recognizes the long history of admissibility of such evidence, but as the *Daubert* Court stated, "Hypotheses . . . that are incorrect will eventually be shown to be so."

15. These histories are drawn largely from my introductions to the forensic science specialties discussed in DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS & JOSEPH SANDERS, *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* (1997).

admissibility of scientific evidence.

## I. A Brief History of Legal Tests for the Admissibility of Scientific Evidence

Most modern discussions of the admissibility of scientific evidence have remarkably little to say about the period before *Frye v. United States*,<sup>16</sup> as if courts before 1923 subjected expert witnesses to no test.<sup>17</sup> But the first reported decision affirming the propriety of the use of such "skilled witnesses"<sup>18</sup> was in 1782, in the case of *Folkes v. Chadd*.<sup>19</sup> There are earlier recorded uses, however. One such early recorded use occurred in 1699 in the notorious murder trial of Spencer Cowper,<sup>20</sup> though the judge in that case clearly did not regard it as a novelty even then.<sup>21</sup>

A review of cases from the decades preceding *Frye* reveals that the courts, at least as far back as the civil war, were using what recently has been termed the "commercial marketplace test."<sup>22</sup> Under the marketplace test, courts sought to determine whether the pro-

---

16. 293 F. 1013 (D.C. Cir. 1923).

17. See, e.g., MCCORMICK ON EVIDENCE § 203 (John W. Strong et al. eds., 4th ed. 1992) ("[A] special rule for scientific evidence originated in 1923 in *Frye* . . ."); Moenssens et al., *supra* note 4 ("The oldest of these standards has come to be known as the 'Frye' test . . .").

18. In the 18th century, the term "skilled witness" covered everyone we would today refer to as an expert. In modern times, the term seems to have returned to referring to practical, as contrasted with scientific, experts. See FED. R. EVID. 702 advisory committee's note.

19. 99 Eng. Rep. 589, 590 (K.B. 1782) (approving expert testimony by engineers to elucidate the reason behind dam deterioration).

20. 13 Howell's State Trials 1105 (1699). Spencer Cowper was a barrister charged with the murder of Sarah Stout, a young lady with whose family he regularly stayed when in Hertford attending the assizes. He had been at the Stout residence until late on the evening before her body was discovered in the Priory River. The expert testimony of over a dozen witnesses, including many of the leading English physicians and surgeons of the time, concerned whether a person dead by strangulation before being thrown into the river would float or sink, as opposed to one who had drowned herself in the river to commit suicide. It further concerned the likely cause of certain marks on the body, and how much water might be expected in the lungs in a death by drowning. Testimony was also heard on experimentation involving the drowning of dogs, though the results were inconclusive. Cowper was acquitted. Since he ended his career on the Common Pleas bench nearly thirty years later, he has the distinction of being one of the few English superior court judges ever to have stood trial for common murder.

21. See *id.* at 1123.

22. David L. Faigman et al., *Check Your Crystal Ball at the Courthouse Door, Please: Exploring the Past, Understanding the Present, and Worrying about the Future of Scientific Evidence*, 15 CARDOZO L. REV. 1799, 1803-05, 1805 n.13 (1994). The following discussion of the antecedents leading to *Frye*, and the meaning of *Frye* in light of its antecedents, is drawn from that article.

posed expert witness was offering relevant facts or opinions that were beyond the ken of the fact finder, and whether the proposed witness was "qualified" as an expert. The implicit test of expertise was whether there was a commercial market for the witness's learning. If a person could make a living selling the knowledge at issue, then expertise presumably existed. Although courts sometimes spoke of an expert's "greater study respecting certain subjects,"<sup>23</sup> or having "made the subject upon which he gives his opinion a matter of particular study,"<sup>24</sup> it is clear that some degree of prosperity in the practice of the occupation or profession claiming that knowledge almost always accompanied the expertise.<sup>25</sup> In effect, the marketplace determined whether valid knowledge existed by endowing it with commercial value.

The marketplace test had serious flaws, two of which concern us here. First, the marketplace test was incapable of distinguishing astrophysics from astrology. The market values both of them. Commercial value, then, is not a measure of scientific validity. A second problem was that some fields have little or no life in any commercial marketplace. Indeed, the fields that are the focus of this Article have little or no function outside of their possible courtroom utility. The courtroom is their marketplace.<sup>26</sup>

In light of the test of admissibility that already existed when the *Frye* case arose, we can both readily understand the necessity for inventing the *Frye* test and appreciate the real importance of this minor corollary to the marketplace test.

In proffering an early form of polygraph testing, *Frye* presented the court with an unfamiliar problem. How was the validity of such asserted expertise to be evaluated and a judgment made as to its admissibility? The technique was new. There were no professional polygraph examiners yet, and no developed market for their services. Perhaps the technique was valid, perhaps not. Judge Van Orsdel apparently realized that the conventional legal test offered no solution.<sup>27</sup> The asserted expertise offered in *Frye* demanded of the court a different test than the one that had long served the law.

Judge Van Orsdel found a solution that closely resembled the existing test. The entire opinion took up only two pages of the Fed-

---

23. *Jones v. Tucker*, 41 N.H. 546, 548 (1860).

24. *Id.*

25. See Faigman et al., *supra* note 22, at 1804 & n.15.

26. Use of the term "forensic sciences" is emblematic of this fact. See BLACK'S LAW DICTIONARY 648 (6th ed. 1990) (defining "forensic" as "belonging to courts of justice").

27. Courts in similar situations have been less attentive, and sometimes dealt with this problem by using no test. See *infra* Part III. Moreover, Judge Van Orsdel himself did not bother to use the test when its implications would have been inconvenient. See *infra* note 33 and accompanying text.

eral Reporter. The critical language is:

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.<sup>28</sup>

At bottom, this is nothing more than the familiar market test relocated to a different marketplace. Where there is no commercial market, the intellectual or professional marketplace could serve instead. The *Frye* test is still a marketplace test, the real evaluation still is conducted outside of the court and outside of the law, and it still is incapable of distinguishing astrology from astrophysics.

But the *Frye* corollary did three new and important things. First, the alternative marketplace allows knowledge to be evaluated even if the knowledge is too new to be marketed commercially, or if there is no hope of ever marketing it commercially. This is important to forensic identification science, which has little or no existence outside of the investigation and prosecution of crimes.<sup>29</sup> The traditional commercial marketplace test, if used today, likely would exclude all of the subfields addressed in this article.<sup>30</sup>

Second, the *Frye* test separated the expertise from the expert, thereby creating explicit legal recognition of the notion that a body of asserted knowledge has an existence separate from any individual, and it is that body of asserted knowledge that has to be evaluated apart from any particular individuals who might seek to bring it to court.

Third, and no doubt inadvertently, *Frye* replaced buyers with sellers as the real judges of the validity of the offerings. The commercial marketplace test, even with its serious weaknesses, had the virtue of allowing buyers to assess the value of purported expertise and whether it was, "therefore," valid. Under the *Frye* variant, that control was transferred to the people who produced the knowledge and offered it (and themselves) to the courts.<sup>31</sup>

---

28. *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

29. If crime laboratories, typically attached to police departments, did not employ them, most of the experts whose work is discussed in this Article would be unemployed. After all, what activities of life other than detective work make use of fingerprints, tool-marks, bitemarks, and so on?

30. Ironically, it would admit polygraph experts because, despite their frequent exclusion from the courtroom, they enjoy considerable employment in business settings, government agencies, and police investigations.

31. Thus, the *Frye* test, far more than its commercial marketplace predecessor, is



If the first myth of the *Frye* test is that it was a revolutionary judicial invention, the second myth is that it thereafter dominated the courts' scientific admissibility decisions.<sup>32</sup> In actuality, the *Frye* corollary was so minor a variant that it went unnoticed for decades. Indeed, Judge Van Orsdel himself ignored it in another landmark scientific evidence case he handed down on the very same day he issued the *Frye* opinion.<sup>33</sup> *Frye* was not cited by a single other court, federal or state, for a decade. During the first quarter century after its publication, *Frye* was cited in only eight federal cases and five state cases.<sup>34</sup> That amounts to one case every other year in the entire country. During its second quarter century, it was cited fifty-four times in federal cases and twenty-nine times in state cases.<sup>35</sup> The *Frye* test really was "discovered" only in the past few decades. Consequently, in the 1980's *Frye* was being cited as much each year as it had been in its first fifty years added together.<sup>36</sup>

The Federal Rules of Evidence, adopted in 1975, rejected the *Frye* corollary in favor of a test that focused on the scientific validity of the proffered expert testimony. Of course, we did not learn this about the Federal Rules until the Supreme Court's unanimous decision in *Daubert v. Merrell Dow* in 1993. Until that time, most of the federal circuits<sup>37</sup> and half the states<sup>38</sup> incorporated *Frye* into Rule

---

vulnerable to the widely varying standards of the producers of knowledge rather than the standards of the consumers. As a result, under the *Frye* corollary, because the courts must rely on the standards set within each field, they will find themselves more readily accepting the offerings of less rigorous fields and less readily accepting the offerings of the more rigorous fields.

32. See, e.g., PETER HUBER, *GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM* 14, 176 (1991) ("Thereafter, federal courts, widely copied by the states, were bound by the *Frye* rule . . ." and "[f]rom 1923 until the mid-1970's, the *Frye* rule made some attempt to hold expert witnesses to similar standards . . .")

33. *Laney v. United States*, 54 F. 412, 416 (D.C. Cir. 1923) (admitting firearms identification expert testimony).

34. See Faigman et al., *supra* note 22, at 1808 n.25.

35. See *id.*

36. See *id.*

37. Seven circuits had interpreted the Federal Rules of Evidence to embody *Frye*: *Christophersen v. Allied-Signal Corp.*, 939 F.2d 1106, 1110 (5th Cir. 1991) (per curiam); *United States v. Two Bulls*, 918 F.2d 56, 60 (8th Cir.) *reh'g granted, vacated en banc, and remanded*, 925 F.2d 1127 (8th Cir. 1991); *United States v. Gillespie*, 852 F.2d 475, 480 (9th Cir. 1988); *Kropinski v. World Plan Executive Council—U.S.*, 853 F.2d 948, 956 (D.C. Cir. 1988); *United States v. Metzger*, 778 F.2d 1195, 1203 (6th Cir. 1985); *United Ellis v. Int'l Playtex, Inc.*, 745 F.2d 292, 303-04 (4th Cir. 1984); *States v. Smith*, 869 F.2d 348, 350 (7th Cir. 1989). Two other circuits had held that the *Frye* test was incompatible with the Federal Rules: *United States v. Jakobetz*, 955 F.2d 786, 794 (2d Cir.), *cert. denied*, 113 S. Ct. 104 (1992); *United States v. Downing*, 753 F.2d 1224, 1237 (3d Cir. 1985).

38. See Thomas Bohan & Erik Heels, *The Case Against Daubert: The New Scientific Evidence "Standard" and the Standards of the Several States*, 40 J. FORENSIC SCI. 1030, 1036 tbl.2 (1995).

702—despite the total absence of any language in the Rules or the Advisory Committee Notes invoking *Frye* or the concept of general acceptance.<sup>39</sup>

The test embodied in FED. R. EVID. 702, according to *Daubert*, requires the trial court to conduct a preliminary hearing<sup>40</sup> in which it decides whether “the reasoning or methodology underlying the testimony is scientifically valid and properly can be applied to the facts at issue.”<sup>41</sup> The application of that test requires a court to determine scientific validity with the help of several non-exclusive guidelines. Table 1 lists the essential language of the first three of those guidelines in the order in which they appear in the opinion, juxtaposed against the first three major sections of conventional scientific journal articles. The parallelism is striking. This is, of course, entirely consistent with *Daubert*’s goal: to direct the attention of courts (away from concern with the popularity of a belief among scientists) to a scientific proposition’s underlying evidence and logic.

**Table 1                      The Daubert “Factors”**

Major Daubert Factors	Components of a Scientific Journal Article
1. Testability (falsifiability): “whether [the subject matter] can be (and has been) tested”	Introduction
2. Good methodology: “Peer review and publication . . .” “submission to the scrutiny of the scientific community . . .” “increases the likelihood that substantive flaws in methodology will be detected”	Methods
3. “known or potential rate of error”	Results

Perhaps the most hazardous myth of the *Frye* test is that it is a more stringent, less liberal, standard than that of other tests,<sup>42</sup> notably *Daubert*.<sup>43</sup> The two tests focus on different attributes of purported

39. “Frye’s ‘general acceptance’ test was superseded by the Rules’ adoption. The Rules occupy the field . . .” *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 587 (1993). Further: “Nothing in the text of this Rule establishes ‘general acceptance’ as an absolute prerequisite to admissibility. Nor does respondent present any clear indication that Rule 702 or the Rules as a whole were intended to incorporate a general acceptance standard. The drafting history makes no mention of Frye . . .” *Id.* at 588.

40. *Id.* at 592-93 (citing FED. R. EVID. 104(a)).

41. *Id.*

42. See Huber, *supra* note 32, at 198-205 (arguing that the rigors of the *Frye* test are necessary to keep “junk science” out of the courtroom).

43. The *Daubert* opinion itself shares the illusion that *Frye* is necessarily the more conservative test:

scientific knowledge. Whether a proffered scientific theory or technique passes muster is going to depend upon how it fares on the attribute to which the test directs a judge's attention. Figure 1 will help clarify this analysis and its implications.

**Figure 1**                    **Is Daubert a More Rigorous or a More Relaxed Test of Admissibility than Frye?**

FRYE: General Acceptance	DAUBERT: Scientific Foundation	
	Strong	Weak
High	Both Admit	Frye admits Daubert excludes
Low	Frye excludes Daubert admits	Both Exclude

Any given theory or technique can be based on a strong scientific foundation or a weak one and, independently, it can enjoy "general acceptance in the particular field in which it belongs" or not. A theory or technique that is based upon valid science *and* enjoys general acceptance would be admitted by either test. A theory or technique that is not based upon valid science *and* does not enjoy general acceptance would be excluded by either test. The interesting situations are where the two attributes are discordant for any given type of scientific evidence. Where proffered knowledge is based on a solid scientific foundation but has not yet gained general acceptance within its field, *Frye* would exclude but *Daubert* would admit. This is the situation that is usually envisioned when the two tests are compared, and so *Frye* is seen as the more conservative test. But where proffered knowledge has only a weak scientific foundation and yet enjoys general acceptance within its field, the *Frye* test will admit but the *Daubert* test would exclude. In this situation the *Frye* test is not conservative at all, but downright radical. This fourth situation occurs more frequently than most commentators have been aware. Among those fields where this particular discordance exists, *Daubert*

---

Nothing in the Rules as a whole or in the text and drafting history of Rule 702, which specifically governs expert testimony, gives any indication that "general acceptance" is a necessary precondition to the admissibility of scientific evidence. Moreover, *such a rigid standard would be at odds with the Rules' liberal thrust and their general approach of relaxing the traditional barriers to "opinion" testimony.*

*Daubert*, 509 U.S. at 588 (emphasis added). This has also been the view after *Daubert*. See, e.g., *Joiner v. General Electric Co.*, 78 F.3d. 524 (11th Cir. 1996) (opining that *Daubert* intended to make it easier, not harder, to present conflicting scientific views to a jury).

has provoked deep concern. The forensic identification sciences are among those concerned fields, and not without reason, as the decisions in *United States v. Starzecpyzel*<sup>44</sup> and *Williamson v. Reynolds*<sup>45</sup> demonstrate. In short, *Daubert* subjects the forensic identification sciences to more rigorous scrutiny than *Frye* had, and the field's practitioners know it.

We should not leave this subject without a word about the implications for appellate review of *Daubert's* interpretation of the Federal Rules' admissibility provisions. It has been common for appellate courts to state that they review the admissibility of scientific evidence under a deferential standard.<sup>46</sup> At such moments, courts appear to believe that the decision at issue is merely a determination of adjudicative fact. On other occasions, appellate courts seem to realize that a ruling on admissibility, on whether an asserted expertise exists or not, is really a holding of law.<sup>47</sup> Accordingly, they do not refrain from looking beyond the record, from doing their own research, or from deciding *de novo* whether the trial court was correct in its assessment of the science.<sup>48</sup> Moreover, lower courts customarily defer to decisions of admissibility when they are made by higher courts, treating them as if they were law.<sup>49</sup>

As a matter of logic, the latter view seems to be the more sensible. If each trial court may decide for itself whether an expertise exists, then appellate courts would have to uphold contradictory conclusions of different trial courts.<sup>50</sup> By treating scientific evidence

---

44. 880 F. Supp. 1027, 1047 (S.D.N.Y. 1995) (declaring that handwriting experts do not pass *Daubert*). But see discussion of this case and its progeny, *infra* Part III.A.

45. 904 F. Supp. 1529, 1558 (E.D. Okla. 1995) (stating: "This court has been unsuccessful in its attempts to locate any indication that expert hair comparison testimony meets any of the requirements of *Daubert*. Not even the 'general acceptance' standard is met.").

46. See, e.g., *Spencer v. Commonwealth*, 393 S.E.2d 609, 621 (Va. 1990), *cert. denied*, 498 U.S. 908 (1990) ("If there is a conflict, and the trial court's finding is supported by credible evidence, it will not be disturbed on appeal."). The U.S. Supreme Court has recently adopted this view for the federal courts. See *General Elec. Co. v. Joiner*, 118 S. Ct. 512 (1997).

47. Other questions, such as an expert's qualification, are more universally viewed as matters of fact to be reviewed on a clear error standard.

48. See, e.g., *Commonwealth v. Curnin*, 565 N.E.2d 440, 443 (Mass. 1991) ("In making the determination whether [the scientific] test is generally accepted, courts may properly consider not only the evidence in the record but also the reasoning and conclusions of other courts and the writings of experts. In these circumstances, an appellate court makes its own determination without regard to the conclusions of the trial or motion judge." (citations and footnotes omitted).)

49. See, e.g., many of the cases of voiceprint admissibility cited in Appendix.

50. See *Dunagin v. City of Oxford, Miss.*, 718 F.2d 738, 748-49 n.8 (5th Cir. 1983), (concluding that when two district courts reach contrary conclusions on the same empirical, legislative fact issue, the responsibility passes to the court above to decide, *de novo*,

admissibility decisions as holdings of law, appellate courts can bring about more consistent treatment of such evidence as well as increase the accuracy of decisions by increasing judicial scrutiny.

The Supreme Court's choice to place these decisions under Federal Rule of Evidence 104(a) (rather than 104(b)) may provide a further suggestion that these decisions are a matter of law and not of fact, since the decision is completely removed from the traditional trial fact finder, the jury.

In a slightly different context, these issues have been explored in some detail by Professors Monahan and Walker. The essential legal analysis is this: When reviewing *adjudicative* (case-specific) facts, appellate courts can safely defer to the discretion of the trial court, and employ the clearly-erroneous standard, because those facts are unique to the case at bar and the trial court is in the best position to scrutinize the evidence.<sup>51</sup> But rulings on scientific evidence embody findings of *legislative* (trans-case) facts.<sup>52</sup> They inherently apply to cases beyond the case at bar.<sup>53</sup> And their truth or falsity has nothing to do with the credibility of witnesses or anything else that a trial court is in a better position to assess. It stands or falls on the body of scientific knowledge, whose literature can be evaluated equally well, if not better, by the appellate court. If one trial court holds that people have unique, identifying voiceprints and that a technology exists that can detect those differences, it is incoherent to allow the courtroom next door to hold the opposite.

## II. A Brief Intellectual History of Forensic Identification Science

This Part reviews the task of forensic identification science, the field's origins, the evidence on which its claims rest, and possible reasons for its arrested development as a science.<sup>54</sup>

---

which is correct).

51. See Laurens Walker & John Monahan, *Social Facts: Scientific Methodology as Legal Precedent*, 76 CAL. L. REV. 877, 881 (1988).

52. See *id.*

53. See John Monahan & Laurens Walker, *Social Authority: Obtaining, Evaluating and Establishing Social Science in Law*, 134 U. PA. L. REV. 477 (1986) (arguing that social science research is more analogous to a rule of law than a fact when used to formulate a rule of law). See also Laurens Walker & John Monahan, *Social Frameworks: A New Use of Social Science in Law*, 73 VA. L. REV. 559 (1987) (analyzing a large set of cases of "framework" uses of scientific evidence, wherein the court decides the general phenomenon as a matter of law, and the factfinder applies the general knowledge to decide the particularized factual question at issue).

54. For a broad overview of the intellectual history of forensic science, the interested reader also will want to consult the works of Thorwald, *infra* note 74, Thornton, *infra* notes 55 and 72, and other works cited *infra*.

## A. The Problem of Identification

Forensic identification science has selected for itself—or had thrust upon it—a project that is unknown to other fields: the unique identification or, more properly, *individualization* of various objects, including persons, distinct from all others in the world. “Criminalistics is the science of individualization.”<sup>55</sup>

The question posed is whether a bullet can be traced back to the one and only one barrel through which it was fired, a signature to the hand that wrote it, a bite mark to the mouth of the biter, cut bolts to the instrument that cut them, and so on. Affirmative answers are offered daily in courtrooms across the country as firearms examiners, document examiners, forensic odontologists, tool mark experts, and numerous other forensic identification scientists purport to identify *the* gun, hand, mouth, tool, and so on, that left its traces at a crime scene.<sup>56</sup> DNA typing is merely the latest addition to the family of forensic identification sciences. Each member of this family subscribes to the same assumptions and draws its inferences from the same basic logic. Typically, testimony based on such identifications is offered to place a defendant at the scene of a crime.

The capacity to make such identifications depends on the validity of a series of premises: That many kinds of biological and physical entities exist in unique, one-of-a-kind form; that they leave correspondingly unique traces of themselves; and that the techniques of observation, measurement, and inference employed by forensic science are adequate to link these traces back to the object that produced them. The claim usually has been presented in essentially this strong form: Individualization is “absolute specificity and absolute identification.”<sup>57</sup>

---

55. James W. Osterburg, *The Evaluation of Physical Evidence in Criminalistics: Subjective or Objective Process?*, 60 J. CRIM. L. & CRIMINOLOGY 97, 97 (1969); see also John Thornton, *The General Assumptions and Rationale of Forensic Identification*, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 20 (David L. Faigman et al. eds., 1997).

56. Thus, forensic identification science should be distinguished from the forensic applications of normal science. The forensic sciences can be divided into two branches: the categorization-and-quantitation branch and the individuation branch. The first branch looks like normal applied science: It borrows the principles of chemistry or biology or some other established field of normal science, and with that it seeks to categorize a substance, object, or event and, if possible, to measure its quantity. It answers such questions as: Is there alcohol in the corpse's blood and, if so, how much? The second branch is quite different. Its basis in an established conventional science is limited or non-existent. There is no university or industrial department of fingerprints or toolmarks or handwriting. The focus of this article is on this latter branch—the “science” of individuation.

57. David A. Stoney, *What Made Us Ever Think We Could Individualize Using Sta-*

The goal of individualization contrasts with conventional science of virtually every kind. "Individualization is unique to forensic science."<sup>58</sup> Normal science is concerned with grouping objects and events into meaningful classes, discovering systematic relationships among these classes, and developing and testing theoretical explanations for those shared attributes and relationships. While normal science looks only *between* classes, forensic identification science purposefully ignores "class characteristics" and looks *within* classes. While normal science is concerned with establishing regularities, forensic science is concerned with exploiting irregularities among objects within classes. Its central assumption is that objects possess enough differences that on adequate inspection one object cannot be mistaken for another.

## B. The Theoretical and Empirical Basis of Individualization

The question arises as to the basis for believing that all things are unique and that individualization is possible. Many forensic scientists are content to assert that no two of various types of objects can be alike, and leave it at that. For example, one textbook asserts, without support: "It has been shown empirically, with theoretical support, that fingerprints are unique. No two persons . . . possess identical ridge characteristics."<sup>59</sup> Thoughtful efforts to justify these claims usually begin with notions from probability theory, but when those scholars realize that probability theory simply cannot get there from here, they look in vain for another route.<sup>60</sup>

For example, after using a manifestly probabilistic thought experiment to defend the proposition that no two fingerprints can be alike, Cummins and Midlo conclude: "It is unfortunate that this approach carries the implication that a complete correspondence of two patterns might occur, when as a matter of fact . . . such duplication is beyond the range of possibility."<sup>61</sup> Failing to make the case with their

---

tistics, 31 J. FORENSIC SCI. SOC'Y. 197, 197 (1991). With the advent of DNA typing, ironically, forensic identification science has begun to be recognized as a probabilistic and less than completely certain endeavor. See Michael J. Saks & Jonathan J. Koehler, *What DNA "Fingerprinting" Can Teach the Law About the Rest of Forensic Science*, 13 CARDOZO L. REV. 361 (1991). Similarly, the law's discussions of science once implied the expectation science of a kind of certainty that real science could fairly be said to specialize in *not* offering, a view which the law, in *Daubert*, has begun to abandon. See Edward J. Imwinkelried, *Evidence Law Visits Jurassic Park: The Far-reaching Implications of the Daubert Court's Recognition of the Uncertainty of the Scientific Enterprise*, 81 IOWA L. REV. 55, 64-71 (1995).

58. PETER R. DE FOREST ET AL., FORENSIC SCIENCE 7 (1983).

59. *Id.* at 332.

60. See Stoney, *supra* note 57.

61. HAROLD CUMMINS & CHARLES MIDLO, FINGER PRINTS, PALMS AND SOLES:

best rationale, they retreat to anecdotes, assumptions ("nature never repeats"),<sup>62</sup> and appeals to intuition ("common sense rejects as fantastic the idea" of two being alike).<sup>63</sup> A small but perhaps growing number of forensic identification scientists accepts the unavoidable: such identifications are in reality estimates of probability.<sup>64</sup>

Unfortunately, the probabilities employed by traditional forensic identification science are subjective and intuitive. Only the newest of these, DNA typing, takes the burdens of the probabilistic nature of forensic identification science seriously.<sup>65</sup>

Many forensic scientists who concede the inherently probabilistic nature of their enterprise nevertheless refrained from undertaking data collection and the calculation of empirically based probabilities: "In most [handwriting identification] problems it would be impossible, or at least extremely impractical, to measure mathematically the degree of probability of accidental coincidence. . . . These basic conditions prevent arithmetic determination of a probability factor."<sup>66</sup> Others, however, have been distressed by the "almost complete lack of factual and statistical data pertaining to the problem of establishing identity"<sup>67</sup> in their areas, and have started the belated work of building a rigorous foundation for forensic identification science.<sup>68</sup>

---

AN INTRODUCTION TO DERMATOGLYPHICS 154 (1943).

62. *Id.* at 150.

63. *Id.* at 153-54.

64. See THE USE OF STATISTICS IN FORENSIC SCIENCE (D.A. Stoney & C.G.G. Aitken eds., 1991); Alfred A. Biasotti, *The Principles of Evidence Evaluation as Applied to Firearms and Toolmark Identification*, 9 J. FORENSIC SCI. 428 (1964); Alwyn Cole, *The Search for Certainty and the Uses of Probability*, 25 J. FORENSIC SCI. 826-33 (1980); Daniel Van Gelder, *An Examination of the Application of Bayes' Theorem to Trace Evidence Evaluation—The Criminalist's Perspective*, 11 J. POLICE SCI. & ADMIN. 377 (1983); S. Keiser-Nielsen, *Dental Identification: Certainty v. Probability*, 9 FORENSIC SCI. 87 (1977); K.W. Kemp et al., *Some Statistical Aspects of Similarity and Identification*, 14 FORENSIC SCI. SOC'Y J. 335 (1974); P.L. Kirk & C.R. Kingston, *Evidence Evaluation and Problems in General Criminalistics*, 9 J. FORENSIC SCI. 434 (1964).

65. The problem for forensic identification sciences other than DNA is that one cannot claim unique individualization and prove that with probability theory and data. If one relies on probabilities, one has to admit, as did Cummins and Midlo and others, that there is a measurable probability of coincidental matches. Then one must collect the data necessary to make the calculations and base one's testimony on those calculations. All of that is hard work. Furthermore, the rewards for that hard work are to have to temper one's claims from unique individuality and virtual flawlessness to (merely) a low probability of error, and to have courts regard one's offerings with greater skepticism. See Saks and Koehler, *supra* note 57 (predicting that the probabilistic and data-based approach of DNA typing will become the norm for all of forensic identification science).

66. ORDWAY HILTON, *SCIENTIFIC EXAMINATION OF QUESTIONED DOCUMENTS* 9 (rev. ed. 1982).

67. Alfred A. Biasotti, *A Statistical Study of Individual Characteristics of Fired Bullets*, 4 J. FORENSIC SCI. 34, 34 (1959).

68. See P.D. Barnett & R.R. Ogle, *Probabilities and Human Hair Comparison*, 27 J.



But such people are forensic science pioneers; the community itself has yet to settle the scientific ground they long ago staked their claim to.

No articulated theory exists that explains why unique identifiability must be the order of the universe.<sup>69</sup> The origins of the forensic scientist's notion that no-two-are-alike can be traced to Adolph Quetelet, the 19th Century Belgian statistician and sociologist better remembered today as the father of descriptive social statistics.<sup>70</sup> Based on a statistical concept we will examine shortly, Quetelet hypothesized that nature never creates biological duplicates. Alphonse Bertillon, a ne'er-do-well lad whose influential father obtained a position for him as a minor clerk in the Paris Prefecture of Police, had learned of Quetelet's hypothesis from his father and grandfather, respected practitioners of medicine, statistics, anthropology, and demography. With his father's help, in the early 1880s Bertillon overcame the resistance of his superiors and used Quetelet's hypothesis to develop the first system of forensic identification, known as anthropometry, or *bertillonage*. Bertillon measured eleven different physical features of each prisoner and assembled the measures into special files. If Quetelet were right, Bertillon would be able to identify prisoners who had been arrested before and on re-arrest were using aliases. At the same time, Bertillon had a source of data with which to begin testing Quetelet's hypothesis.<sup>71</sup>

With the invention of *anthropometry*, forensic identification science was born. Bertillon's system proved useful in detecting recidivists. An interesting question is, what ended *bertillonage*? Forensic scientists usually recount that what doomed *bertillonage* was the discovery of several cases of different prisoners with indistinguishable Bertillon measurements combined with the advent of a more accurate system, namely, fingerprinting.<sup>72</sup> But the historical evidence indicates that *bertillonage* lived on, and was even introduced into new

---

FORENSIC SCI. 272 (1982); Biasotti, *supra* note 64; B.D. Gaudette & E.S. Keeping, *An Attempt at Determining Probabilities in Human Scalp Hair Comparison*, 19 J. FORENSIC SCI. 599 (1974); Stoney, *supra* note 57; David A. Stoney & John I. Thornton, *A Critical Analysis of Quantitative Fingerprint Individuality Models*, 31 J. FORENSIC SCI. 1187 (1986); Van Gelder, *supra* note 64.

69. See James W. Osterburg, *What Problems Must Criminalistics Solve?*, 59 J. CRIM. L., CRIMINOLOGY & POLICE SCI. 427, 429 (1968) (contrasting forensic science's individualization with other fields' practice of merely situating a sample in a class).

70. See 9 THE NEW ENCYCLOPEDIA BRITANNICA 855 (15th ed. 1995).

71. See HENRY T. F. RHODES, ALPHONSE BERTILLON: FATHER OF SCIENTIFIC DETECTION (1956).

72. See John Thornton, *Criminalistics—Past, Present, Future*, 11 LEX ET SCIENTIA, 1, 14 n.53 (1975). In the instance of handwriting identification, the falsification of the fundamental belief that no two signatures are alike did not give the least pause to forensic document examiners. See *infra* notes 94-95 and accompanying text.

law enforcement agencies well after the advent of fingerprinting, being still in use in parts of the United States as late as the 1930s. *Bertillonage* was, after all, better established and rested on at least as sound a scientific foundation. Moreover, the stories of the falsification of *bertillonage* appear to have been fabricated by supporters of fingerprinting.<sup>73</sup> Fingerprinting did, of course, eventually replace *anthropometry*, no doubt because fingerprinting was more convenient and efficient, and because people often left their prints, but never their Bertillon measurements, at crime scenes.<sup>74</sup>

Bertillon nevertheless had established the concept that on some attributes people varied to such a degree that certain measures could be useful in identifying them. Once the courts accepted fingerprint evidence as being uniquely identifying, fingerprints became an icon for every other kind of identification evidence—tool marks, bullets, bite marks, handwriting, voiceprints, shoe prints, broken glass, and so on. That DNA typing often is referred to as “genetic fingerprinting” and voice spectrography as “voiceprints” illustrates the lineage of identification science. Examiners would explain that their object of study was just like fingerprints, and courts usually believed them.<sup>75</sup>

It mattered not that Quetelet’s hypothesis had not been proved. It mattered not that the hypothesis never could be proved. (Asserting that no two fingerprints are alike is analogous to claiming that all

---

73. See David L. Grieve, *The Identification Process: Traditions in Training*, 40 J. FORENSIC IDENTIFICATION 195 (1990) (debunking the story that *bertillonage* erroneously identified Will West).

By contrast, consider that there is much better evidence of the falsification of this fundamental tenet of handwriting identification, that no two people’s signatures are indistinguishably alike, and yet that finding has had no impact on the beliefs or practices of document examiners. See *infra* notes 94 and 140 and accompanying text.

74. Though an enemy of fingerprinting, Bertillon was interested in other evolving forensic sciences in addition to anthropometry. He testified as a handwriting expert in the infamous Dreyfus affair, erroneously identifying Lt. Dreyfus as the author of an incriminating letter, leading to Dreyfus’s conviction for treason. The actual author and traitor later confessed. The wrongful conviction of Dreyfus was a major political embarrassment to the French Third Republic. On his deathbed, for his lifetime contributions to French and world police work, Bertillon was offered the highest medal the French government had to offer, if only he would publicly concede the error of his testimony in the Dreyfus case. Bertillon refused. See JURGEN THORWALD, *THE CENTURY OF THE DETECTIVE* 89-90 (1965).

75. For example, the first appellate court to admit toolmark evidence opined: Courts are no longer skeptical that by the aid of scientific appliances the identity of a person may be established by finger prints. There is no difference in principle in . . . [determining] that the same tool that made one impression is the same instrument that made another impression. The edge on one blade differs as greatly from the edge on another blade as the lines on one human hand differ from the lines on another.

State v. Clark, 287 P. 18, 20 (Wa. 1930).

swans are white. "No matter how many instances of white swans we may have observed, this does not justify the conclusion that *all* swans are white."<sup>76</sup>) Nor did it matter that the step from natural objects<sup>77</sup> (e.g., anthropometry and fingerprints) to manufactured ones (e.g., tools, guns) went beyond anything that Quetelet had proposed or Bertillon had thought to test. What for the social statisticians of the 19th Century had been a hypothesis to be tested became for most forensic scientists and courts of the 20th Century an article of faith, past questioning and beyond question.

### C. The Multiplication Rule and Beyond

Asked for hard evidence—even today—each of the subfields of forensic identification science rests its claims of infinite variation and unique identifiability on nothing more than what Quetelet had offered, namely the multiplication rule of probability applied to populations. The essential idea of this concept is that if objects vary on a number of independent (i.e., uncorrelated) dimensions, the probability of occurrence of any one combination of characteristics is found by multiplying together the probabilities associated with each dimension. Such calculations typically produce probabilities that are vanishingly small. Having made this general point, the next step in the argument—and it is offered by the progenitors of each forensic identification science subfield—is to appeal to the audience's intuition to make the leap into concluding that no two handwritings,<sup>78</sup> tool marks,<sup>79</sup> fingerprints,<sup>80</sup> gun barrels,<sup>81</sup> or whatever, could be alike.

Heavy reliance on the multiplication rule as the foundation of individualization encounters numerous problems.<sup>82</sup> First, of course,

---

76. KARL R. POPPER, *THE LOGIC OF SCIENTIFIC DISCOVERY* 27 (1959). Moreover, failing to search with vigor for white swans—or for violations of the assumption that no two are alike—does not protect a field's claims, it weakens them.

77. Remember, it was "nature never repeats." See CUMMINS & MIDLO, *supra* note 61.

78. See ALBERT S. OSBORN, *QUESTIONED DOCUMENTS* 219 (2d ed. 1929) (arguing that variations in individuals' handwriting "cannot be exactly duplicated in the writing of any other person").

79. See Luke S. May, *The Identification of Knives, Tools and Instruments, A Positive Science*, 1 AM. J. POLICE SCI. 246, 247 (1930) (stating that identifying marks left by tools lead to "obvious" identification).

80. See V. Balthazard, *No Two Finger Prints Alike*, SCIENTIFIC AMERICAN, Aug. 18, 1911, at 166.

81. See Calvin H. Goddard, *Scientific Identification of Firearms and Bullets*, 8 J. AM. INST. CRIM. L. & CRIMINOLOGY 254 (1926) (comparing linear striations formed on bullets fired from a gun to a "fingerprint of that particular barrel").

82. See Stoney, *supra* note 57.

as Cummins and Midlo<sup>83</sup> realized, probabilistic models cannot prove absolutes, such as that no two are alike. This is not physics, where two objects cannot occupy the same place at the same time. This is micro-taxonomy, where no law of nature prevents two or many objects from falling into the same category. Any given bridge hand has a probability of occurrence of less than 1 in 600 billion. Yet it would be obvious nonsense to presume that nature has arranged the universe so that once a bridge hand is dealt it will never be dealt again.<sup>84</sup> Nevertheless, in forensic science there has been a leap from notions of probability to belief in a doctrine of unique individuality. Even if unique individuality did rule the universe, establishing the validity of a forensic technique would require testing the system of measurement and classification as well, even (or especially) if its principal tool is human perception and judgment.<sup>85</sup>

Second, many of the rule's applications to forensic science may violate the independence assumption, so that the probabilities are not so small as the usual illustrations imagine them to be.<sup>86</sup>

Third, with the exception of recent work involving biological markers, such as DNA typing, the various forensic identification sciences have not taken the trouble to collect data on populations of forensically relevant objects so that the probability of erroneous

---

83. See CUMMINS & MIDLO, *supra* note 61, at 154 ("Under the circumstances it is impossible to offer decisive *proof* that no two fingerprints bear identical patterns, but the facts in hand demonstrate the soundness of the working principle that *prints from two different fingers are never identical*.").

84. For example, consider the following from Marcia Clark in State's Closing Argument at 21, *People v. Simpson*, No. BA097211, 1995 WL 697928 (Cal. Super. Ct. 1995):

[L]adies and gentlemen, his blood on the rear gate with that match that makes him one in 57 billion people that could have left that blood, I mean there is what, five million [sic] people on the planet, that means you would have to go through 57 billion people to find the DNA profile that matches Mr. Simpson's. There is [sic] only five billion people on the planet. Ladies and gentlemen, that is an identification, okay, that proves it is his blood. Nobody else's on the planet; no one.

85. Forensic identification examiners use the naked eye or magnification along with their judgment to decide when two things are alike or different. See David Stoney, *Fingerprint Identification*, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 21-2.1.2(3)(c) (David L. Faigman et al. eds., 1997) ("In fingerprint comparison judgments of correspondence and the assessment of differences are wholly subjective; there are no objective criteria for determining when a difference may be explainable or not.").

86. See FRANCIS GALTON, *FINGER PRINTS* (De Capo ed., 1965) (1892). Galton was well aware of the problem of conditional probabilities: "It is hateful to blunder in calculations of adverse chances, by overlooking correlations between variables, and to falsely assume them independent, with the result that inflated estimates are made which require to be proportionately reduced." *Id.* at 109.

matches<sup>87</sup> can be calculated.<sup>88</sup> Instead, examiners implicitly assume the odds are one-to-infinity.<sup>89</sup>

Fourth, the steps from observation of similarity to the conclusions that are offered to courts must traverse a minefield of potential errors of probabilistic inference that few forensic scientists, and even fewer lawyers or judges, are equipped to navigate.<sup>90</sup>

The preceding problems have led at least two state supreme courts to hold probability-based evidence insufficient to support a criminal conviction.<sup>91</sup> The logic of the identification evidence offered and rejected in those cases is the same as the logic of forensic science. Presumably, if those courts excluded other identification evidence for these shortcomings, they would be constrained to exclude forensic science for the same deficiencies.

Additional reasons exist for preferring forensic identification techniques to be constructed on a foundation of real data and formal probability models. One is that it is the main road from subjective

---

87. While forensic scientists prefer to use the phrase "share a common origin," *see, e.g.* DeForest et al., *supra* note 58, at 51, when opining that a known and a questioned sample cannot be distinguished, I will use the term most familiar to laypersons, namely, "match." This is not to say that it is not important to distinguish among the terms and the definitions they denote. Different forensic scientists prefer different terms and sometimes the same term means different things to different forensic scientists. For example, forensic dentists understand the term "match" to be a "nonspecific term indicating some degree of concordance . . . an expression of similarity without stating degree of probability or specificity." *MANUAL OF FORENSIC ODONTOLOGY* 350 (C. Michael Bowers & Gary L. Bell eds., 3d ed. 1995).

88. Doing so makes shortcomings more apparent, as is evident in the debate over the adequacy of DNA typing data: sample sizes, inclusion of relevant subpopulations, etc. *See* COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, *DNA TECHNOLOGY IN FORENSIC SCIENCE* 9-15 (1992). For most other forensic sciences, these issues are barely addressed or utterly ignored.

89. *See generally* Saks & Koehler, *supra* note 54 (suggesting that the data-based and probabilistic approach of DNA typing will become the norm for all forensic identification sciences).

90. For illuminating discussions of these problems, *see* *STATISTICS AND THE LAW* (Morris H. DeGroot, Stephen E. Fienberg & Joseph B. Kadane eds., 1986); *MICHAEL O. FINKELSTEIN & B. LEVIN, STATISTICS FOR LAWYERS* (1990); David H. Kaye, *The Probability of an Ultimate Issue: The Strange Case of Paternity Testing*, 75 *IOWA L. REV.* 75 (1989); Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 *JUDICATURE* 222 (1993); William C. Thompson & Simon Ford, *The Meaning of a Match: Sources of Ambiguity in the Interpretation of DNA Prints*, in *FORENSIC DNA TECHNOLOGY* 93 (Mark A. Farley & James J. Harrington eds., 1991).

91. *See generally* *People v. Collins*, 438 P.2d 33, 41 (Cal. 1968) (holding that prosecution's offer of a formula in statistical probability constitutes prejudicial error in a criminal case due to absence of data as foundation for the frequency estimates, lack of evidence of independence among attributes, and incorrect interpretation of the resulting probability calculation); *State v. Sneed*, 414 P.2d 858, 862 (N.M. 1966) (holding that mathematical odds were not admissible as evidence to identify defendant, where odds were based on estimates which had not been shown to be valid).

impressions to science.<sup>92</sup> Another is that for several identification techniques, the assumption of no-two-alike has already been empirically disconfirmed. As a result, the need to make more accurate estimations of the reduction in uncertainty afforded by these techniques has become patent. Earlier we noted the possibility of men with identical Bertillon measures. Significantly, *bertillonage* was rare among the forensic identification sciences in that it allowed the possibility of encountering false matches in the course of routine work, because only this technique kept such detailed records and filed each card in a precise position. Under the theory, two cards could not occupy the same space.<sup>93</sup>

For every other identification technique, the only way to find false matches would be to conduct special studies to look for them. I know of only one such effort, by a document examiner who went looking for signatures from different people that were indistinguishably alike. He found them in abundance.<sup>94</sup> In so doing, he fatally falsified the core claim of handwriting identification.<sup>95</sup> Of course, the best way to avoid finding duplicates is not to look for them. As long as one refrains from looking for black swans one's belief that all swans are white is insulated from falsification. More importantly, no one really knows how diagnostic each identification technique is. Only by carrying out appropriate studies could that become known.

In addition, proficiency studies, undertaken only in the past 20 years, revealed varying rates of error.<sup>96</sup> For example, for DNA typing, the error rate has been about 1-2%;<sup>97</sup> for tool marks as much as

---

92. See Keiser-Nielsen, *supra* note 64, at 89; Kirk & Kingston, *supra* note 64, at 435; Osterburg, *supra* note 55, at 101.

93. Fingerprinting is the only other forensic science that keeps a record of each object of interest. But unlike Bertillon records, fingerprint cards are filed by classification, not by singular location. See MOENSSENS ET AL., *supra* note 4, at 503-04, 508-09. Thus, although fingerprints have the potential to make erroneous matches manifest, only with the advent of computerized searching may that potential be realized. All other forensic sciences typically compare a questioned sample to a known one, an examiner decides whether they appear to be the same or different, and these samples are set aside, perhaps never to be looked at again.

94. See John Harris, *How Much Do People Write Alike: A Study of Signatures*, 48 J. CRIM. L. & CRIMINOLOGY & POL. SCI. 647 (1958) (noting the indistinguishable similarity of many signatures).

95. Though neither document examiners nor the courts have taken any notice.

96. See JOSEPH L. PETERSON, ELLEN L. FABRICANT & KENNETH S. FIELD, CRIME LABORATORY PROFICIENCY TESTING RESEARCH PROGRAM 25 (1978); Joseph L. Peterson & Penelope N. Markham, *Crime Laboratory Proficiency Testing Results, 1978-1991, II: Resolving Questions of Common Origin*, 40 J. FORENSIC SCI. 1009 (1995).

97. California Association of Crime Laboratory Directors (CACLD, 1987 and 1988); Koehler, *supra* note 90, at 229.

35%;<sup>98</sup> for handwriting identification about 36%.<sup>99</sup> While these proficiency values are open to alternative calculations and interpretations,<sup>100</sup> what is most significant is that these error rates exist at all. Even if forensic metaphysicians were right, that no two of anything are alike, for fact finders in earthly cases, the problem is to assess the risk of error whatever its source, be that in the basic theory or in the error rates associated with human examiners or their apparatus. That information can best be evaluated and supplied by placing the testimony in a probabilistic context that combines proficiency data, what is known about population base rates, and whatever else is relevant to assessing the probativeness of the testimony.<sup>101</sup>

#### D. A Science Constructed in the Image of the Criminal Law

Forensic science plainly has something of value to offer criminal investigators and the courts. Why, then, does so much of it cling, instead, to an untenable absolutism and committed subjectivity? By contrast, conventional science would have proceeded along a different course, one guided by the necessity of collecting and analyzing data to test assumptions. In court, conventional scientists might be expected to share with the fact finder the analytic basis of their opinions, their data, and their data-based assessments of the risk of error. In short, conventional scientists would collect better data and offer them to the courts without overselling them.<sup>102</sup> Why doesn't forensic science proceed along that more recognizably scientific path?

---

98. Based on my tabulation of the 1980 through 1987 toolmark proficiency testing reports from the Collaborative Testing Service and the Forensic Sciences Foundation.

99. See D. Michael Risinger, Mark P. Denbeaux & Michael J. Saks, *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification "Expertise,"* 137 U. PA. L. REV. 731 (1989). In one of the proficiency tests reported, examiners were in complete agreement with each other (100% reliability) but all were incorrect (0% validity). See *id.* at 745. For further discussion of this issue, see D. Michael Risinger & Michael J. Saks, *Science and Nonsense in the Courts: Daubert Meets Handwriting Identification Expertise*, 82 IOWA. L. REV. 21, 43-44 n.104 (1996).

100. For example, how representative are the tests of the population of identification problems encountered by examiners? Are unrealistically easy or difficult tests included? Are higher rates of error found in blind testing than when examiners know they are being tested? See Lamotte et al., *Comparison of Laboratory Performance with Blind and Mail-Distributed Proficiency Testing Samples*, 92 PUB. HEALTH REP. 554 (1977). Are erroneous exclusions more common than erroneous inclusions? What do subtests reveal? For example, in my tabulations of toolmark proficiency studies, accuracy differed between cutting tools (38% correct) and prying tools (87% correct) and between exemplars supplied by the test manufacturer (98% correct) and those prepared by the technician as part of the test (40% correct).

101. See Saks & Koehler, *supra* note 57, at 369 (supporting the use of proficiency tests to estimate the error rate reported in court testimony).

102. See Commentary, *Fingerprint Trials*, 339 NATURE 491 (1989).

The answer likely is that forensic science grew up in the criminal law. The exigencies imposed on it by police and prosecutors molded it into its contemporary shape.<sup>103</sup> A particularly dramatic demonstration of this is the lengths to which some forensic scientists have been willing to go to provide courts with the testimony prosecutors wanted courts to hear, regardless of the truth. Paul Giannelli has summarized an array of fraudulent science, faked tests, and perjured testimony.<sup>104</sup> But one need not look to such scandalous examples to find the influence of the adversary process at work. Consider the following demands under which forensic science has been required to operate.

In order to win a conviction, the prosecution must prove its case "beyond a reasonable doubt." If the forensic scientist testifies: "I cannot tell these questioned and known evidence items apart, so they probably share a common origin, but of course this is only a subjective estimation based on intuition, because we've never mapped the distribution of what is out there" or "based on our sampling of the population we calculate the probability of a coincidental match to be at the following level of probability," room is left for some doubt. But doubt vanishes if the forensic scientist can say something along these lines: "Because the questioned and the known look alike, and because each person's or object's marks are unique in all the world, I can state with certainty my opinion that the defendant left the markers found at the crime scene."

Gaps will be seized upon by the defense, which will argue to the jury: "If there is even *one* other match out there, that makes two people who might have done it, only one of whom is my client; that implies a 50:50 chance that someone other than my client is guilty. Surely you cannot regard that as guilt beyond a reasonable doubt." Courts have reversed convictions on the reasoning that a merely rare

---

103. Had the reverse been true, had forensic science come into being in the service of the defense, it no doubt would be distorted in the opposite direction, with an exaggerated emphasis on exclusion. Rather than quieting doubts, it would be the master of raising them. Both are perfectly sensible inferences that can be drawn. The emphasis comes not from the science, but from its use by lawyers.

104. See Paul Giannelli, *The Abuse of Scientific Evidence in Criminal Cases*, 4 VA. J. SOC. POL'Y & L. 439 (1997) (proposing that forensic scientists be freed from the pressures that lead to fraud and perjury by removing crime laboratories from the control of police departments). To this extent, then, we are making the same point: That in important respects forensic science is a product of the social arrangements within which it is required to operate and evolve. See generally David Johnston & Andrew C. Revkin, *Report Finds F.B.I. Lab Slipping From Pinnacle of Crime Fighting*, N.Y. TIMES, Jan. 29, 1997, at A1, B8. ("Scientists at the laboratory said they were often stifled in an operation run by nontechnical field agents who had little knowledge of science and who regularly altered reports to help prosecutors.")



probability is not sufficient to prove guilt.<sup>105</sup> From a litigator's viewpoint, there is a world of difference between saying this is *a* match and saying this is *the* match. What modern scientists do as a matter of course—measure the risk of error—presents serious problems for prosecutors and therefore for forensic scientists.<sup>106</sup>

Because of its institutional position within the legal system, the forensic identification sciences have taken on a shape that resembles no other science. Consider these special attributes: No other fields are as closely affiliated with a single side of litigation as forensic science is to criminal prosecution.<sup>107</sup> Police crime laboratories were not begun in order to provide science for police and courts, but as a public relations device.<sup>108</sup> Even today, few of the personnel of crime laboratories have scientific training beyond the undergraduate level, and some not even that.<sup>109</sup> Crime laboratories generate very little research, which to a scientist means they are not doing science, and to a lawyer should say at least that little progress is being made. At best, they apply science, but even that often is not the case.<sup>110</sup> Progress might come from their colleagues in industrial or academic departments. But there are no industrial uses of what forensic identification scientists do. And the number of university programs to train

---

105. See, e.g., *State v. Sneed*, 414 P.2d 858, 861-62 (N.M. 1966).

106. The continued pressure created by this legal goalpost may be illustrated by an advertisement by Genetic Design, a DNA typing laboratory: "When there's no room for doubt. Let Genetic Design Perform Your Forensic DNA Analysis and Put an End to Uncertainty." Program of the American Academy of Forensic Sciences, 45th Annual Meeting, February 15-20, 1993, Boston, MA, outside rear cover. This is particularly ironic, given that DNA typing is the prototypical forensic identification science that has revealed to courts the need for sampling and probability calculations in forensic identification. See Saks & Koehler, *supra* note 57.

107. Most of the fields we are talking about in this article have no function except as part of the government's investigation and prosecution of crime. Forensic science use of biological markers provides the exception that tests this rule. Unlike toolmarks, bitemarks, handwriting, and so on, which essentially were developed within and for police work, techniques using biological markers were borrowed from disciplines that had a separate and established existence. The approach used by those latter techniques, even in their forensic applications, bears a strong resemblance to their root science, with a concern for empirical data, quantification, and the measurement of error.

108. See Thornton, *supra* note 72, at 7.

109. This is not a matter of academic snobbery. It is in Ph.D. programs where students are trained to "think like a scientist." An analogy might be law firms hiring paralegals to work as attorneys, rather than law school graduates. They can follow a variety of recipes, but as a group their understanding penetrates only so far. In other fields, such as radar or sonar or in hospital laboratories, we call such people technicians. That crime laboratory technicians have come to be called forensic *scientists* is, very likely, a product of the needs of lawyers.

110. Where is the "application" from basic science to forensic identification science? What are the basic sciences from which come fingerprint identification, handwriting identification, and the others? There are none.

forensic scientists can be counted on one's fingers.<sup>111</sup> The maldistribution of forensic scientists so favors the prosecution that the defense has little access to any, which prevents the adversary process from working, as intended, to expose error.<sup>112</sup> The institutional setting of forensic science promotes habits of thought that more closely resemble the thinking of litigators than of scientists. While science pursues knowledge through disconfirmation,<sup>113</sup> prosecutions are won by confirmatory proofs. This confirmatory bias dominates the thinking of most forensic scientists.<sup>114</sup> Where science advances by open discussion and debate, forensic science has been infected by the litigator's preference for secrecy.<sup>115</sup> Tests of the proficiency of crime

---

111. Indeed, for the working forensic scientist, there is an implicit disincentive to ask fundamental questions and make fundamental advances. Fundamental new discoveries risk raising judicial doubts about all that had gone before, and what future research might reveal about past conclusions. No advances means fewer doubts.

The problem is illustrated by an exchange between a judge at a continuing judicial education program and a physician who spoke to the group. The physician showed slides of certain marks on a child's vagina, and explained that those once were thought to indicate sexual abuse of the child, but that subsequent research had shown that they were not. With apparent distress, a judge asked the physician whether she recalled appearing in his court some years before and testifying that such marks *did* support abuse charges, and resulted in the conviction and imprisonment of a defendant. The judge was upset, because the physician's lecture was tantamount to recanting her earlier testimony, and implied that an innocent man may have gone to prison. The physician asked the judge if he would prefer that the medical profession stop advancing its knowledge so that they and he would never have to learn about their past mistakes. This incident took place at the University of Iowa College of Law, involving Iowa judges and a physician from the University of Iowa Hospital. It was related to me by my colleague, Professor Sheldon Kurtz, who was present for the doctor's presentation and the informal exchange following it.

112. Almost uniquely, forensic science finds itself without any regular source of competition to challenge it, to compel it to strengthen and improve its knowledge base, to perform in a first rate fashion. As a result, little progress is made and avoidable errors go unchecked.

113. See Karl Popper, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC KNOWLEDGE 37 (5th ed. 1989) ("[T]he criterion of the scientific status of a theory is its falsifiability, or refutability, or testability"), quoted in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593 (1993) (emphasis deleted).

114. Conventional scientists are said to proceed by falsification, or disconfirmation. That is, they subject hypotheses to tests capable of disproving their validity. See *id.* Crime investigators, including forensic scientists, pursue investigatory strategies that more closely resemble the confirmatory strategies of people in ordinary life. See W.B. Swann, Jr. & T. Giuliano, *Confirmatory Search Strategies in Social Interaction: How, When, Why, and with What Consequences?*, 5 J. SOC. & CLIN. PSYCH. 511 (1987). That is, they try to obtain evidence consistent with the hypothesis of interest. For crime investigators, that means seeking evidence consistent with a suspect's guilt. See, e.g., the description of fingerprint identification given in PAUL L. KIRK, *CRIME INVESTIGATION: PHYSICAL EVIDENCE AND THE POLICE LABORATORY* 71 (2nd ed. 1974).

115. Although in individual cases the release of information is regulated by law, see FED. R. CRIM. P. 16, general knowledge also is kept under wraps. The director of one laboratory attended a defense lawyer's continuing education seminar "undercover" and

laboratories are conducted anonymously, kept secret, and are not routinely published.<sup>116</sup> It is ironic that while studies of the effectiveness and accuracy of so many professional enterprises are available in published literature, the same is not true of a field whose sole purpose is to do some of the public's most public business.

In short, courts and lawyers and the criminal justice establishment within which the forensic identification sciences exist are in all probability the major cause of the arrested development of forensic identification science. The norms of science and of scientific institutions have been too faint and distant an influence on forensic science.

### III. A Series of Encounters

This Part examines the law's leading encounters with several major forensic identifications sciences, encounters which not only set the precedents for the courts' reception of each particular forensic science subfield, but set the patterns for dealing with claims of science in other forensic fields. Lessons emerge from each of these experiences.

#### A. Handwriting Identification: Heads, the Proponent Wins; Tails, the Opponent Loses

Because the ancient and recent history of asserted handwriting identification expertise, and the law's response to it, have been discussed at length elsewhere,<sup>117</sup> this section will provide only a brief summary of that long history.

Handwriting identification expertise is the oldest "forensic science," in the sense of a skill whose only value is in resolving legal disputes. Modern handwriting identification experts claim, in parallel with other forensic identification experts, that writing varies infinitely among people, therefore that no two people write alike,<sup>118</sup> and that they

---

later "noted with alarm" that D.E.A. and F.B.I. manuals were being offered for sale to defense attorneys. MOENSSSENS ET AL., *supra* note 4, at 27-28 n.9. Another example is the fact that proficiency testing results are not published and access to that data is limited. Compare this with the norms of science. See SHARING RESEARCH DATA (S. Fienberg, M. Martin & M. Straf eds., 1985).

116. See PETERSON ET AL., *supra* note 96, at 7 (noting crime laboratories' refusal to participate in proficiency testing unless anonymity and confidentiality were assured); see also Andre A. Moenssens, *Novel Scientific Evidence in Criminal Cases: Some Words of Caution*, 84 J. CRIM. L. & CRIMINOLOGY 1 (1993) (discussing proficiency testing, human error, fraudulent evidence, reliability, and similar issues).

117. See Risinger et al., *supra* note 99, at 751-71. See generally Risinger & Saks, *supra* note 99. Moreover, I want to credit Professor Risinger as a co-author of this subsection of the present article on handwriting identification.

118. Complicating handwriting identification is the additional problem of variation within a writer's work as well as among writers. This factor reduces the likely validity of the claim and increases the risk of error in any given case. Also, see Harris's data tending to

have a special ability to determine the authorship of a piece of writing, including detecting forgeries.

The notion that handwriting can be used to identify an author is very old.<sup>119</sup> The notion that a person can learn to make such an identification by study has also been around for a long time. Attempts to develop a system of such expertise appear to have started in Italy and France in the seventeenth century,<sup>120</sup> and by 1737 were well enough accepted in France to have been incorporated into the law.<sup>121</sup> A century after its introduction in French courts, handwriting expertise entered Anglo-American courts, although the courts welcomed it hesitantly, sometimes admitting,<sup>122</sup> but more often excluding<sup>123</sup> or placing limitations on what the expert could do.<sup>124</sup> Not until an 1854 statute<sup>125</sup> was construed to authorize it, did handwriting identification expertise become admissible in English courts.<sup>126</sup>

In the United States, until passage of the English statute, most American jurisdictions followed English practice and rejected such expertise, though there were some important exceptions. In 1836, with *Moody v. Rowell*,<sup>127</sup> Massachusetts became the first common law jurisdiction to authorize the use of a handwriting expert, reasoning that because the conventional non-expert methods of identifying disputed writings were so poor, the expert was unlikely to be any worse.<sup>128</sup>

While by 1900 a substantial majority of American jurisdictions ac-

---

contradict the claim, *supra* note 94, at 647.

119. "Just as all men do not have the same speech sounds, neither do they have the same handwriting," attributed to Aristotle in HUNTINGDON HARTFORD, YOU ARE WHAT YOU WRITE 43 (1973).

120. The earliest treatise in this line, of a definite graphological cast, appears to be Camillo Baldi, *Trattato come da una lettera missiva si conoscano la natura, e qualità dello scrittore* (Milano, Geo. Batt Bidelli, 1625) [An Essay on the Means of Examining the Character and Qualities of a Writer from His Letters].

121. See Risinger & Saks, *supra* note 99, at 22 n.8.

122. See, e.g., *Goodtitle d. Revett v. Braham*, 4 Term Rep. 497, 498, 100 Eng. Rep. 1139, 1140 (1792).

123. See, e.g., *Carey v. Pitt*, 170 Eng. Rep. 219 (1793) (holding that hand-writing cannot be proved by a person who has never corresponded with him or seen him write); *Stranger v. Searle*, 170 Eng. Rep. 265 (1793) (holding that evidence of hand-writing from comparison of hands not admissible).

124. See *The King v. Cator*, 170 Eng. Rep. 661, 670-71 (C.P. 1802); *Doe d. Mudd v. Suckermore*, 111 Eng. Rep. 1331, 1333, 1344, 1347 (K.B. 1836).

125. See Common Law Procedure Act, 1854, 5 HALSBURY'S STATUTES 340-42 (4th ed. 1993).

126. See Risinger et al., *supra* note 99, at 757-58 n.116.

127. 34 Mass (17 Pick.) 490, 498 (1836).

128. The *Moody* court concluded that "this species of evidence, though generally very slight, and often wholly immaterial, is competent evidence." *Id.* at 498. This seems to be the dominant rationale for the allowance of such testimony in those states which followed Massachusetts' lead over the next fifty to seventy-five years. See Risinger et al., *supra* note 99, at 736-37, nn.21-22 and accompanying text.

cepted such testimony,<sup>129</sup> the prevailing attitude may be best exemplified by the opinion of the New York Court of Appeals in *Hoag v. Wright*,<sup>130</sup> where the court said:

The opinions of experts upon handwriting, who testify from comparison only, are regarded by the courts as of uncertain value, because in so many cases where such evidence is received witnesses of equal honesty, intelligence and experience reach conclusions not only diametrically opposite, but always in favor of the party who called them.<sup>131</sup>

As much or more than most forensic sciences, handwriting expertise was a product of the needs of litigators. When lawyers wanted experts to identify handwriting, and when the courts finally admitted such expertise, there were no handwriting experts. Lawyers seeking to utilize such testimony had to proffer various witnesses who were willing to assert a kind of ad hoc expertise acquired as a side effect of being something else, such as a postal inspector or a bank teller. No practicing forensic document examiner today would concede any expertise to such witnesses.<sup>132</sup> When the legal system agreed to accept such testimony, however, it created a demand which was to be met by people who turned their entire attention to filling it.

Two events finally gained respectability for handwriting expertise. One was the publication in 1910 of Albert S. Osborn's *Questioned Documents*, with an introduction by John Henry Wigmore.<sup>133</sup> Osborn's book, Osborn's personality, and Osborn's friendship with Wigmore, are the cornerstones upon which respect for asserted handwriting identification expertise in the United States was built. Osborn's book set out the theory and practice of the claimed expertise so comprehensively that it is fair to say that all treatments of the subject since have simply been rearrangements or expansions of Osborn's 1910 book.<sup>134</sup> Together, Osborn and Wigmore conducted a quarter century public relations campaign on behalf of "scientific" handwriting identification expertise as practiced by Osborn and described in his book.

The "arrival" of document examination was finally secured by the Lindbergh Baby kidnapping case, *State v. Hauptmann*,<sup>135</sup> in 1935. Os-

---

129. The earliest authority for the admission of such testimony in each United States jurisdiction is set out chronologically in Risinger et al., *supra* note 99, at 788-92, app.3.

130. 66 N.E. 579 (1903).

131. *Id.* at 581; *accord* Miles v. Loomis, 75 N.Y. 88 (1878); Mutual Benefit Life Ins. Co. v. Brown, 30 N.J. Eq. 193 (1878); In re Fuller's Estate, 70 A. 105 (1908).

132. See, e.g., Osborn, *supra* note 78, at 286-88.

133. ALBERT S. OSBORN, QUESTIONED DOCUMENTS (1910).

134. This includes Osborn's own 1929 second edition, *supra* note 78, which had surprisingly little new information on handwriting identification theory or practice. Most of its material on those topics is taken verbatim from the 1910 edition, *supra* note 133.

135. 180 A. 809 (N.J. 1935).

born was the chief witness called to testify that Bruno Richard Hauptmann had written all of the ransom notes found or sent after the abduction of the son of Charles A. Lindbergh.<sup>136</sup> The public seemed to need to believe Hauptmann was guilty, wanted him convicted, and was grateful to those who supplied the evidence. Osborn became a celebrity. For 60 years following the affirmance of *State v. Hauptmann*, no reported opinion rejected handwriting expertise, nor was much skepticism displayed towards it. Rather, it became universally accepted as scientific and dependable.<sup>137</sup>

After standing unquestioned for most of this century, a re-evaluation of handwriting identification expertise has resulted from the Supreme Court's decision in *Daubert*. To date, three federal decisions have undertaken this new look, and each has concluded that forensic document examination lacks a scientific basis.<sup>138</sup>

In *United States v. Starzecpyzel*,<sup>139</sup> Judge McKenna examines at length the claims of handwriting identification expertise to scientific status, and rejects them.<sup>140</sup> Because asserted handwriting identification

---

136. See *id.* at 822.

137. This may be best illustrated by the following quotation from *In re Estate of Sylvestri*, a New York decision contrasting sharply with the criticism found in *Hoag*: "Since that rather cynical observation was made by our highest court in *Hoag*, examiners of questioned documents, as handwriting experts prefer to be called, have attained more respectable standing in the courtroom." 55 A.D. 2d 916, 390 N.Y.S.2d 598 (App. Div. 1977). As a New Jersey court observed in *Morrone v. Morrone*, 130 A.2d 396 (N.J. 1957), after the *Hauptmann* case, handwriting identification expertise could no longer be regarded as "the lowest order of evidence and . . . accorded little evidential weight." *Id.* at 400. Osborn recognized the centrality of the *Hauptmann* case. Concerning it, he wrote in 1940, "[i]t can be correctly stated that in that little one hundred year old courtroom at Flemington, N.J., the scientific examination and proof of the facts in document cases was nationally recognized and firmly established as a New Profession." Albert S. Osborn, *A New Profession*, 24 J. AM. JUD. SOC. 303 (1940), reprinted in Osborn, *supra* note 78, at 311. For a similar evaluation, see JAMES V.P. CONWAY, *EVIDENTIAL DOCUMENTS* 210 (1959).

138. See *United States v. Velasquez*, 64 F.3d 844, 850 (3rd Cir. 1995) (likening handwriting identification to other non-sciences, but claiming nonetheless to apply the *Daubert* test "in an exercise of caution"); *United States v. Starzecpyzel*, 880 F. Supp. 1027, 1029-41 (S.D.N.Y. 1995) (holding that forensic document examination expertise is outside the scope of *Daubert* but may be admissible as nonscientific "skilled" testimony); *United States v. Ruth*, 42 M.J. 730, 732 (Army Ct. Crim. App. 1995) (essentially adopting Judge McKenna's opinion in *Starzecpyzel* as its own).

139. 880 F. Supp. 1027 (S.D.N.Y. 1995). The author, called as a witness by the opponent of admissibility, testified at the *Daubert* hearing concerning the methodological and statistical requirements of empirical verification generally, and the lack of such studies or other data testing the claims of handwriting identification expertise.

140. The defense presented evidence demonstrating that: After a century of existence, the field of forensic handwriting identification had produced little or no research demonstrating the truth of any of its major claims, such as about the nature of handwriting or document examiners' special skills, that some research showed the claim of unique variation in handwriting to be false by finding numerous signature doubles (*supra* note 94), and that

expertise is not scientific, however, he reasons that it need not pass *Daubert's* validation requirements, because those apply only to sciences. He then analogizes handwriting examiners to harbor pilots, who learn to do something dependably by experience.<sup>141</sup> And he shifts the burden of proof on the issue of whether a non-science expertise exists from the proponent to the opponent.<sup>142</sup> Nevertheless, Judge McKenna then declared himself persuaded that the inferences as to genuineness of the signature at issue in the case before him "can be performed with sufficient reliability to merit admission."<sup>143</sup>

*United States v. Velasquez*<sup>144</sup> is somewhat more problematic than *Starzecpyzel* because it is more ambiguous. The *Velasquez* court never clearly explained whether it was judging the admissibility of handwriting identification expertise under the *Daubert* criteria for scientific evidence, or under some other standard for "technical or other specialized knowledge."<sup>145</sup> It cites *Starzecpyzel* as authority for the proposition that the criteria for judging admission of "scientific expertise set out in *Daubert* are inapplicable to non-scientific handwriting identifica-

---

proficiency studies of handwriting examiners showed high error rates. The court concluded: "Were the Court to apply *Daubert* to the proffered FDE [forensic document examiners] testimony, it would have to be excluded. This conclusion derives from a straightforward analysis of the suggested *Daubert* factors—testability and known error rate, peer review and publication, and general acceptance—in light of the evidence adduced at the *Daubert* hearing." *Starzecpyzel*, 880 F. Supp. at 1036. "In sum, the testimony at the *Daubert* hearing firmly established that forensic document examination, despite the existence of a certification program, professional journals and other trappings of science, cannot, after *Daubert*, be regarded as 'scientific . . . knowledge.'"). *Id.* at 1038.

141. The difference is that harbor pilots learn from the instant feedback they receive if they ground a ship or deliver it to the wrong wharf. But forensic identification, including handwriting, is an altogether different task. Neither handwriting examiners themselves, nor lay judges or jurors, can tell whether the identification is correct or not. Whereas the criterion of success or failure in the world of harbor pilots is immediately obvious to all, the accuracy of the conclusions of document examiners remains opaque to all, including them.

142. According to Judge McKenna, the defense "presented no evidence, beyond the bald assertions [of its experts], that FDEs [forensic document examiners] cannot reliably perform this task. Defendants have simply challenged the FDE community to prove that this task can be done reliably. Such a demonstration of proof, which may be appropriate for a scientific expert witness, has never been imposed on 'skilled' experts." *Starzecpyzel*, 880 F. Supp. at 1046. The implication that the ultimate risk of non-persuasion as to reliability is ever on the opponent of a proffer of evidence is remarkable, both as a general notion of evidence law and more specifically in the context of *Daubert*. Under Federal Rule of Evidence 104(a), the proponent of expert evidence must show by a preponderance that the evidence is admissible. See *Daubert v. Merrell Dow*, 509 U.S. 579, 593 n.10 (1993) (citing *Bourjaily v. United States*, 483 U.S. 171 (1987)).

143. *Starzecpyzel*, 880 F. Supp. at 1046. Judge McKenna went on to fashion a jury instruction to be given in advance of the expert's testimony to explain that the testimony was not the result of a scientific process, so that the jurors would have no misconceptions in that regard. See *id.* app. at 1050-51.

144. 64 F.3d 844 (3rd Cir. 1995).

145. FED. R. EVID. 702.

tion.”<sup>146</sup> The *Velasquez* court then goes on to say, however, that “[i]n an exercise of caution” it will review the proffered expertise “for qualifications of reliability and fit as those factors have been explicated in *Daubert*,”<sup>147</sup>—as though the proffered document examination expertise might be a science. But the court then proceeds never even to mention the validity criteria for scientific evidence actually set out in *Daubert*. It concludes by saying, “we agree with the district court that Ms. Bonjour’s proposed testimony concerned ‘scientific, technical or other specialized testimony’ and was sufficiently reliable to be admissible.”<sup>148</sup> This conclusion touches all possible bases for admitting expert evidence under Federal Rule of Evidence 702, and commits to none. At best it is an implied *Starzecpyzel* analysis, and not by any means a *Daubert* scientific validity analysis.

One exception to this pattern has emerged. In the trial of Timothy McVeigh in the Oklahoma Bombing case,<sup>149</sup> the court prohibited document examiners from testifying to any identification of the author of a writing unless the proponent could convince the court at a *Daubert* hearing that their claims rested on a scientific foundation.<sup>150</sup> The government declined to attempt to do so and no expert document examiners testified at the trial.

This oldest of forensic sciences, originally admitted because it was “unlikely to be any worse” than lay testimony, now provides us with the first major preview of how courts under *Daubert* may come to analyze not only asserted handwriting identification expertise, but any of the major forensic identification sciences which, upon examination, are found to contain little science. The problem with the *Starzecpyzel* analysis is that once the court deprives an area of the label “science,” the validation standards for admissibility may plummet almost to zero. Anomalously, the burden shifts to the opponent to prove affirmatively that the experts cannot do what they claim they can do.

The treatment of handwriting examiners highlights the inadequacy of *Daubert* in addressing only the “scientific” prong of Federal Rule of Evidence 702. Under *Starzecpyzel*, once a proffered scientific expertise fails, it simply is evaluated for admission as a non-scientific expertise. What test exists to guide courts in assessing the validity of non-scientific expertise, the “technical or other specialized knowledge,” prongs of Federal Rule of Evidence 702? At present, none. *Star-*

---

146. *Velasquez*, 64 F.3d at 850.

147. *Id.*

148. *Id.* at 851.

149. See *United States v. McVeigh*, No. 96-CR-68, 1997 WL 47724 (D. Colo. Feb. 5, 1997) (ruling on pre-trial motion).

150. Judge Matsch’s ruling still would have allowed general, non-identifying testimony from asserted handwriting experts.



*zecpyzel* teaches us how large a hole that can leave, one so large that it is capable of swallowing *Daubert* in one gulp. In *McVeigh* Judge Matsch found another way to admit a document examiner's testimony while recognizing that the field on which the testimony is based fails the *Daubert* test.

In dealing with non-scientific expertise under Federal Rule of Evidence 702, several considerations should be noted. One is that validity can remain the touchstone. If there are no data confirming the validity of a *field's* claims, then it is not difficult to insist on empirical data demonstrating the special skills of the *particular witness*.<sup>151</sup> Science, real science, simply makes it easier to evaluate the knowledge claims, because real science produces data as a matter of its very nature which can be offered for the court's evaluation.<sup>152</sup> Dealing with non-science will not be so easy. As Professor Imwinkelried has observed in urging courts to develop rational validity standards for non-scientific evidence, "the epistemology of nonscientific expert knowledge is quite different from that of scientific propositions. . . . [T]he development of objective validation standards for nonscientific opinion is likely to prove to be a more difficult task than the formulation of such tests for scientific testimony."<sup>153</sup> While Judge McKenna's approach might be justifiable for a harbor pilot's testimony, in a purely forensic area, such as handwriting identification, some higher standard of affirmative proof would seem to be needed to insure that the conclusions proffered can be arrived at dependably.

## **B. Fingerprints: Vouched for by Dr. Twain and God**

That expert testimony based upon fingerprint evidence to prove identity is admitted in every jurisdiction of the United States<sup>154</sup> is widely known. Less well known is the advent of judicial acceptance of the technique during the second decade of the Twentieth Century. This evidentiary development is characterized by meager judicial scrutiny combined with rapid spread of acceptance among numerous juris-

---

151. For example, suppose a witness claims special ability as a tea taster. It would be relatively easy to construct tests of that asserted skill. Samples of teas of various kinds could be prepared in double-blind fashion and presented to the purported expert who would try to identify and rate the various teas. The responses would be compared to the known tea types, age, chemical composition, or whatever is relevant, and the results of how well the proffered witness performed could be reported to the court for its evaluation.

152. Indeed, the absence of data is the first hint that something is not science.

153. Edward J. Imwinkelried, *The Next Step After Daubert: Developing a Similarly Epistemological Approach to Ensuring the Reliability of Nonscientific Expert Testimony*, 15 CARDOZO L. REV. 2271, 2294 (1994).

154. See Annotation, *Fingerprints, Palmprints, or Bare Footprints as Evidence*, 28 A.L.R. 1115 (1953).

dictions. Such speed is especially surprising considering that fingerprint identification presented the courts with a novel claim (infinite individuality) in an astonishingly strong form (infallibility),<sup>155</sup> and considering the recent failings of anthropometry, the defective first child of forensic individuation, which had made very similar claims.<sup>156</sup>

Case law upholding the admission of fingerprint evidence begins in Illinois in 1911 with *People v. Jennings*.<sup>157</sup> Within the decade New Jersey,<sup>158</sup> New York,<sup>159</sup> Nevada,<sup>160</sup> and Texas<sup>161</sup> joined Illinois in approving the admissibility of fingerprint evidence. These initial jurisdictions established the rationale for admissibility. Little more than the passage of time was necessary for eventual universal acceptance. In the next ten years twelve more states joined. And by the end of the 1930's all but five states were formal members of the club.<sup>162</sup>

These cases, germinal not only for fingerprint identification but for the many other forensic individualization techniques soon to spawn in its path, invested little effort assessing the merits of the proffered scientific evidence. Rather, for the most part, these courts casually cited treatises on criminal investigation, or general approval of science, or, eventually, other cases admitting such evidence.

For example, in *Jennings*, expert opinion based on fingerprints was the principal ground of identification. The court recognized the novelty of the expertise at issue, noting that "the courts of this country do not appear to have had occasion to pass on the question."<sup>163</sup> In upholding the admissibility of fingerprint expertise, the *Jennings* court cited two general encyclopedias,<sup>164</sup> three treatises on crime investigation methods,<sup>165</sup> and one recent English case.<sup>166</sup> Nowhere in the opinion, however, does the court articulate the basis of the expertise it is evaluating, or discuss any scientific evidence bearing on the empirical claims, or illuminate the technique's theoretical premises, or explain

---

155. Recall that once these concepts were accepted on behalf of fingerprint identification, numerous other forensic individualization areas made the same claim by analogizing themselves to fingerprints. See *supra* note 75 and accompanying text.

156. See *supra* notes 70-71 and accompanying text.

157. 96 N.E. 1077 (Ill. 1911).

158. See *State v. Cerciello*, 90 A. 1112, 1115 (N.J. 1914).

159. See *People v. Roach*, 109 N.E. 618 (N.Y. 1915).

160. See *State v. Kuhl*, 175 P. 190, 195-96 (Nev. 1918).

161. See *McGarry v. State*, 200 S.W. 527 (Tex. Crim. App. 1918).

162. The cases are collected at FAIGMAN ET AL., *supra* note 15, § 21-1.0 n.9.

163. *Jennings*, 96 N.E. at 1081.

164. See 10 ENCY. BRITANNICA 376 (11th ed., 1910); 5 NELSON'S ENCYCLOPEDIA 28 (1907).

165. See HANS GROSS, CRIMINAL INVESTIGATION 277 (Adams' transl., 1907); LEONHARD FELIX FULD, POLICE ADMINISTRATION 342 (1909); OSBORN, *supra* note 133, at 479.

166. See *In re Castleton's*, 3 Crim. App. 74 (1909).

why anyone should believe that fingerprint examiners can do what they claim the ability to do. Nor do the cited sources fill that gap.<sup>167</sup> In addition, the court also referred to four experts who testified on behalf of fingerprint identification, each of whom had been studying or practicing fingerprint examination for three to four years before the trial.<sup>168</sup> But the court's opinion shares nothing of what, if anything, these witnesses had to say on the validity of fingerprint identification.

Here is the totality of what the court had to say in its review of the asserted science of fingerprint identification:

These authorities state that this system of identification is of very ancient origin, having been used in Egypt when the impression of the monarch's thumb was used as his sign manual, that it has been used in the courts of India for many years and more recently in the courts of several European countries; that in recent years its use has become very general by the police departments of the large cities of this country and Europe; [of] the great success of the system in England, where it has been used since 1891 in thousands of cases without error . . . .<sup>169</sup>

Based on the preceding, the *Jennings* court concluded:

We are disposed to hold from the evidence of the four witnesses who testified, and from the writings we have referred to on this subject, that there is a scientific basis for the system of finger print identification, and that the courts are justified in admitting this class of evidence . . . .<sup>170</sup>

Since the opinion is without any empirical or theoretical substance, how can the court or a reader of the opinion be persuaded of the existence of the asserted expertise? With equal ease, the court could have satisfied itself as to the validity of astrology.

The second American case to consider the admissibility of fingerprint evidence, *State v. Cerciello*,<sup>171</sup> neither cited nor explained anything

---

167. Nor would one expect them to. With the possible exception of Osborn's treatise on handwriting, *supra* note 133, they do not even purport to be presentations of the basic science underlying techniques of crime investigation.

168. See *People v. Jennings*, 96 N.E. 1077, 1082 (Ill. 1911).

169. *Id.* at 1081. The allusion to "thousands of cases without error" obviously begs the question of validity. In actual disputed cases it rarely, if ever, is possible to tell whether the identification was correct or not; that is why the issue was before the factfinder. This has been a major problem in validating many forensic techniques. See sources cited at note 96, *supra*. For discussions of this problem, see OFFICE OF TECHNOLOGY ASSESSMENT, U.S. CONGRESS, *THE SCIENTIFIC VALIDITY OF POLYGRAPH TESTING: A RESEARCH REVIEW AND EVALUATION* 29-43 (1983); John Thornton, *Courts of Law v. Courts of Science: A Forensic Scientist's Reaction to Daubert*, 1 SHEPARD'S EXPERT AND SCIENTIFIC EVIDENCE QUARTERLY 480 (1994). Take *Jennings* itself as an example: How might one confirm whether fingerprint identification was in error or not? There is no criterion against which to test the correctness of the examiner's conclusion.

170. *Jennings*, 96 N.E. at 1082.

171. 90 A. 1112 (N.J. 1914).

whatsoever concerning the expertise at issue. This court's scientific assessment was nothing more than a generalized endorsement of scientific progress:

[I]ts admission as legal evidence is based upon the theory that the evolution in practical affairs of life, whereby the progressive and scientific tendencies of the age are manifest in every other department of human endeavor, cannot be ignored in legal procedure, but that the law, in its efforts to enforce justice by demonstrating a fact in issue, will allow evidence of those scientific processes which are the work of educated and skillful men in their various departments . . . .<sup>172</sup>

So much for the *Cerciello* court's scientific reasoning. The court's legal reasoning amounted to this: the admission of expert opinion, "one of the prominent exceptions of the general rules of evidence,"<sup>173</sup> was permitted on so many other matters that the court could hardly refuse to permit another exception. Moreover, the jury could be relied upon to give the testimony whatever weight was appropriate.<sup>174</sup>

The *Roach* court did no more than to cite *Castleton* and *Jennings*.<sup>175</sup> Concerning the latter, the *Roach* court commented, credulously: "The opinion of Chief Justice Carter in that case contains an instructive and learned discussion of this whole subject."<sup>176</sup> The opinion offers no citations to any scientific materials or any discussion of the principles claimed to be the foundation of the technique. The court focused on the "qualifications" of the witness rather than the content of the science.<sup>177</sup> The court held: "In view of the progress that has been made by scientific students and those charged with the detection of crime in police departments . . . we cannot rule as a matter of law that such evidence is incompetent."<sup>178</sup>

In Texas, *McGarry v. State*<sup>179</sup> rested its opinion squarely on *Jennings*, literally adopting the Illinois opinion as its own.<sup>180</sup> After quoting at length from *Jennings*, *McGarry* held simply: "We conclude that the evidence of the witness was admissible."<sup>181</sup>

The quality of judicial scrutiny of fingerprint evidence rarely ex-

---

172. *Id.* at 1114.

173. *Id.*

174. By what magic of intuition the jury was to do this we can only guess, since the jury would learn even less about the subject than the court had.

175. See *People v. Roach*, 109 N.E. 618, 623 (N.Y. 1915).

176. *Id.*

177. See *id.* ("Before testifying to his opinion as to the identity . . . the witness explained fully his qualifications, specified the circumstances upon which he predicated his opinion, and swore that he was able to express an opinion with reasonable certainty.").

178. *Id.*

179. 200 S.W. 527 (Tex. Crim. App. 1918).

180. See *id.* at 528-30.

181. *Id.* at 530.

ceeded that of *Jennings*, and sometimes it fell far shorter. While some cases made reference to actual early research studies on fingerprints,<sup>182</sup> others cited Mark Twain's novel, *Pudd'nhead Wilson*,<sup>183</sup> as authority for the infallibility of fingerprint evidence,<sup>184</sup> or appealed to far higher authority:

"God's finger print language," the voiceless speech, and the indelible writing imprinted on the fingers, hand palms, and foot soles of humanity by the All-Wise Creator for some good and useful purpose . . . [namely,] the ultimate elimination of crime . . . [by] unquestionable evidence of identity in all cases.<sup>185</sup>

Before long, courts had ample precedents from sister jurisdictions to cite as authority for the infallibility of fingerprint evidence. Popular and judicial intuitions about fingerprints are so strong that not a case can be found that entertains any serious doubt about the scientific perfection that has been achieved by fingerprint examination.<sup>186</sup>

What is disappointing about the fingerprint admissibility cases is that these courts made no serious, substantive inquiry into the body of knowledge on which they had the responsibility to pass judgment.

Of course, what is found in the opinions has a lot to do with the legal test applied. The reasoning of the earliest cases amounts to: "We're letting so much else in, we might as well let this in, too." Unlike the *Frye* court, these courts did not seem to realize the legal novelty of the situation before them: proffered expert witnesses who had not come from a commercial marketplace in which they had proven themselves. Later cases had the illusory luxury of precedent, reasoning in effect: "Courts in other states are letting in fingerprint evidence, so we can too." The *Frye* test would not have produced much more. The *Frye* test would have required these courts to ask whether people in the field of fingerprint examination believed in fingerprint examination, and they would have found that fingerprint examiners did. The court still would have been looking at something other than the science of

---

182. See, e.g., GALTON, *supra* note 86; SIR WILLIAM J. HERSCHEL, THE ORIGIN OF FINGER PRINTING (Ams Press ed., 1974) (1916); HENRY FAULDS, GUIDE TO FINGER-PRINT IDENTIFICATION (1905).

183. MARK TWAIN, PUDD'NHEAD WILSON (Bantam 1981) (1894).

184. See *State v. Kuhl*, 175 P. 190, 191 (Nev. 1918); *Stacy v. State*, 292 P. 885, 886 (Okla. Crim. App. 1930).

185. *Kuhl*, 175 P. at 191 (quoting approvingly from FREDERIC AUGUSTUS BRAYLEY, BRAYLEY'S ARRANGEMENT OF FINGER PRINTS IDENTIFICATION AND THEIR USES. FOR POLICE DEPARTMENTS, PRISONS, LAWYERS, BANKS, HOMES, TRUST COMPANIES . . . AND IN EVERY BRANCH OF BUSINESS WHERE AN INFALLIBLE SYSTEM OF IDENTIFICATION IS NECESSARY (1909)). In fairness to the *Kuhl* court it must be noted that its opinion was perhaps the most erudite, citing some of the most scholarly works on fingerprints—as well as some of the silliest.

186. See MOENSSENS ET AL., *supra* note 4, at 518 (fingerprint identification is "universally admitted in this country").

the matter.

Certain habits of judicial thought interfered with the courts' inquiries about this, as well as any other, purported science. One is the inclination of a court to be content to cite supportive precedents, rather than to look into the content, the logic, the reasoning of the cited opinions, at least when the judges' initial inclinations appear to be reinforced by the other cases. Where a precedent is a command from a court above, of course, a court below has little choice but to obey. But where a precedent is from a sister state, and therefore has only persuasive value, the cogency of the opinion is everything. The courts of other states that contented themselves with citing *Jennings* and its progeny were being lazy. *Jennings* itself, and other very early cases, made a similar mistake in citing encyclopedias and treatises, as if citing those is the same as citing a case from a court above. Those are not law, at least not unless and until a court makes a finding of legislative fact and thereby makes them an integral part of a holding.<sup>187</sup> Those courts had a duty to make a serious, intellectually defensible, inquiry into the subject matter, a duty they failed to discharge.

A modern court, compelled to apply the conventional scientific criteria required by *Daubert*, would find no help in the earlier cases. Nor would the research literature on fingerprints, nearly a century later, provide the support a modern court would be searching for. Consider what a well respected, scientifically literate, scholar and practitioner of fingerprint examination has written:

Efforts to assess the individuality of DNA typing make an excellent contrast [to fingerprint identification]. There has been [concerning DNA typing] intense debate over which statistical models are to be applied, and how one should quantify increasingly rare events. To many, the absence of adequate statistical modeling, or the controversy regarding calculations, brings the admissibility of the evidence into question. Woe to fingerprint practice were such criteria applied! As noted earlier, about a dozen models for quantification of fingerprint individuality have been proposed. None of these even approaches theoretical adequacy, however, and none has been subjected to empirical validation. Apart from illustration of the intense variability in fingerprint patterns, and the inability of simple minutia counts to quantify this variability, these models *occupy no role* in the routine professional practice of fingerprint examination. Indeed, inasmuch as a statistical method would suggest qualified (non-absolute) opinions, the models are rejected on principle by the fingerprint profession.<sup>188</sup>

Although in principle fingerprint identification depends upon an objective, probabilistic inquiry, its practitioners use no probability

---

187. See *supra* notes 47-53 and accompanying text.

188. Stoney, *supra* note 85, § 21-2.3.1.

models and have no probability data to use. They rely on intuition and assumptions<sup>189</sup> that have not been tested rigorously, and they persist in treating identifications as absolute, not probabilistic.<sup>190</sup>

Fingerprint evidence may present courts applying *Daubert* with their most extreme dilemma. By conventional scientific standards, any serious search for evidence of the validity of fingerprint identification is going to be disappointing. Yet the intuitions that underlie fingerprint examination, and the subjective judgments on which specific case opinions are based, are powerful. When and if a court agrees to revisit the admissibility of fingerprint identification evidence under *Daubert*, the *Daubert* approach—that courts may admit scientific evidence only if it meets contemporary standards of what constitutes science—is likely to meet its most demanding test: A vote to admit fingerprints is a rejection of conventional science as the criterion for admission.<sup>191</sup> A vote for science is a vote to exclude fingerprint expert opinions.

### C. Toolmarks: Reversals without Reasons

Tool mark experts compare markings left at crime scenes by various kinds of tools to similar tools in the possession of suspects in order to try to determine if the suspect's knife or crowbar or wire cutters or whatever left the marks found at the crime scene.

One of the earliest appellate decisions to consider the admissibility of modern tool mark identification expertise was *State v. Fasick*,<sup>192</sup> a Washington State case which, in 1928, rejected the proffered testimony.<sup>193</sup> In that case, a murder had been committed and the body covered with fir branches cut from nearby trees.<sup>194</sup> The government offered Luke S. May, a pioneering forensic scientist who would become one of the founders of tool mark identification.<sup>195</sup> May had made sample cuttings of fir branches with the suspect's knife and examined microscopic marks left by the blade, comparing them with the marks left

---

189. "The criteria for absolute identification of an individual through fingerprint comparison are wholly dependent on the professional judgment of a fingerprint examiner. When a fingerprint examiner determines that there is *enough* corresponding detail to warrant the conclusion of absolute identification, then the criteria have been met." *Id.*

190. "Our standard is a popularly held concept based on subjective criteria and empirically successful practice, not one that was reached though conventional scientific experimentation and statistical evaluation." *Id.*

191. And *Daubert* with it.

192. See 270 P. 123 (Wash. 1928), *aff'd*, 274 P. 712 (Wash. 1929).

193. See *id.* at 124.

194. See *id.*

195. The dust jacket to his semi-autobiographical book on forensic science characterizes May as "America's Sherlock Holmes." LUKE S. MAY, CRIME'S NEMESIS (1936).

in the branches found covering the body.<sup>196</sup> He concluded that the crime scene branches had been cut with the same knife, thus tending to place the defendant at the crime scene.<sup>197</sup> The trial court initially excluded, but later admitted, May's testimony.<sup>198</sup> The Washington Supreme Court, however, found the logic behind the expert's opinion unconvincing. Wrote the Court: "[You] could not tell in a thousand years whether the two pieces were cut by the same knife."<sup>199</sup> On rehearing a year later, the Court affirmed its initial rejection of tool mark identification expertise.<sup>200</sup>

Eighteen months after deciding *Fasick* and only six months after re-affirming itself, in *State v. Clark*,<sup>201</sup> the Washington Supreme Court was presented with remarkably similar evidence, this time in the context of a rape case in which fir boughs and saplings had been cut and used by a rapist to construct a blind from which to attack his victim.<sup>202</sup> Again a knife, again cut fir branches, again Luke May the proffered expert witness.<sup>203</sup> But this time the Court held that expert opinion about whether the defendant's knife cut the branches was admissible: "The photomicrographs . . . conclusively establish, we are convinced, as doubtless the jury were, that the cuts were made with the same blade."<sup>204</sup> Although the *Clark* court superficially distinguishes the *Fasick* case in its opinion, the court failed to explain what changed in its understanding of the scientific claims of tool mark identification, referring only to the larger number of striation comparisons examined in the latter case.

In an article written the same year *Clark* was decided, May discussed the case, describing what he had offered to the court.<sup>205</sup> He asserted that he had "conclusively established" that the same knife was used. His inference of that identification was derived by familiar probabilistic reasoning:

Invoking the law of probabilities, using the algebraic formula for determining combinations and permutations, with only one-third of the marks here shown as factors, there would be only "one" chance of there being another blade exactly like this if every one of the hundred million people in the United States had six hundred and

---

196. See *Fasick*, 270 P. at 124.

197. See *id.*

198. See *id.*

199. *Id.* (quoting the trial court).

200. See *id.*

201. 287 P. 18 (Wash. 1930).

202. See *id.* at 18-19.

203. See *id.*; May, *supra* note 79, at 258.

204. See *Clark*, 287 P. at 20.

205. See May, *supra* note 79, at 253-59.



fifty quadrillion knives each.<sup>206</sup>

Several years later, after praising the *Clark* court for making an "outstanding progressive decision," a "significant step forward,"<sup>207</sup> May described the *Clark* opinion as "a precedent which has already been cited in criminal trials in many other states, thereby legally advancing science in its battle against crime."<sup>208</sup> While it may be that *Clark* was called to the attention of some trial judges, the opinion did not, and perhaps surprisingly so, become the icon for the admission of tool mark expert evidence as, say, *Jennings* had become for fingerprints. *Clark* was cited in only two subsequent appellate opinions, and not for propositions that would have pleased Luke May.<sup>209</sup> Somehow *Clark* went from standing for the "conclusive" power of tool mark identification evidence in 1930 to standing for the admissibility of "a less than exact" process in 1985.<sup>210</sup>

There is little if anything in the *Clark* opinion that could help a judge in a subsequent case to understand why May's claims about a science of tool mark identification were valid. The case merely offered the conclusory and unexplained enthusiasm of a Court that only months before had rendered an opposite opinion on the very same question.<sup>211</sup> Both the exclusion of the purported science in *Fasick* and its admission in *Clark* are unexplained. If the purpose of writing judicial opinions is to explain a court's reasons for reaching its holding, and not merely to enigmatically announce its conclusions, the court failed in both of these cases to meet its obligations.<sup>212</sup> Readers lacking clair-

---

206. *Id.* at 255. This means one chance in  $6.5 \times 10^{25}$  (that is, 65 septillion; 65 followed by 24 zeroes). May does not explain how he made this calculation, but it is worth noting that this is far, far more diagnostic than DNA fingerprinting claims to be.

207. *May, supra* note 195, at 47.

208. *Id.* at 53. Unfortunately, he misses the chance to teach us something about the contrast in the court's thinking from *Fasick* to *Clark* because he does not mention *Fasick* at all.

209. See *Hansel v. Ford Motor Co.*, 473 P.2d 219, 226 (Wash. App. 1970) (a case not involving toolmarks at all); *State v. Bernson*, 700 P.2d 758, 764 (Wash. App. 1985) (standing for the proposition that "[a]n expert's use of 'could have' or 'possibility' has been allowed in other cases where a less than exact scientific process is involved").

210. *Bernson*, 700 P.2d at 764.

211. Yet the *Clark* court was ebullient in its praise of this new specie of evidence:

Courts are no longer skeptical that by the aid of scientific appliances the identity of a person may be established by finger prints. There is no difference in principle in the utilization of the photomicrograph to determine that the same tool that made one impression is the same instrument that made another impression. The edge on one blade differs as greatly from the edge on another blade as the lines on one human hand differ from the lines on another. This is a progressive age. The scientific means afforded should be used to apprehend the criminal.

*State v. Clark*, 287 P. 18, 20 (1930). Nor do the briefs submitted to the Court provide clues as to what changed so dramatically in the court's understanding of toolmark identification from *Fasick* to *Clark*.

212. We can be sure that the changed outcome was not due to changed personnel of the

voyance will never understand why the court believed that tool mark identification expertise did not, and then did, exist.

After so interesting a start in the case law, the subject of the validity of tool mark identification evidence has had surprisingly little appellate exposure in the decades since. From *Fasick* in 1929 to *Ramirez v. State*<sup>213</sup> in 1989, no other cases excluding tool mark identification evidence are to be found.

The first federal court to pass on the question did so in 1978 in *Fletcher v. Lane*.<sup>214</sup> In *Fletcher*, the defendant challenged a tool mark expert's positive identification of a screwdriver found in the defendant's home as the one that made pry marks on the victim's door, to the exclusion of all other screwdrivers in the world.<sup>215</sup> The defense argued that such an assertion was insupportable, or at least unsupported, on scientific grounds, and therefore should not have been admitted. The trial judge converted this challenge from one of scientific validity to one of credibility: "Petitioner's . . . contention is essentially an attack on the credibility of the expert testimony. The credibility of a witness is a matter for jury determination . . ."<sup>216</sup>

An example of the error into which courts may go by so casually scrutinizing expert testimony is provided by *Commonwealth v. Graves*.<sup>217</sup> In this case, expert testimony was offered that scratch marks on the neck of a victim of strangulation matched the defendant's fingernail.<sup>218</sup> The defendant challenged this testimony as lacking sufficient scientific recognition to meet the general acceptance standard.<sup>219</sup> The Pennsylvania court turned away the challenge, concluding:

[T]he methods and techniques used . . . all were consistent with standards of general scientific acceptance in the field of tool-marks, of which, according to the witnesses, testimony as to finger nail wounds is a part.<sup>220</sup>

But consider this comment on the case:

The Pennsylvania reviewing court, however, completely failed to recognize that the class characteristics of the fingernail and the scratch marks, although similar, lacked the necessary individual markings to tie the accused's fingernail to the scratch marks on the

---

court. Three of the votes for exclusion in *Fasick* became votes for admission in *Clark*, including one by the author of *Fasick*.

213. 542 So. 2d 352, 354-55 (Fla. 1989), *rev'd*, 651 So. 2d 1164 (Fla. 1995).

214. 446 F. Supp. 729 (S.D. Ill. 1978).

215. *See id.* at 731.

216. *Id.* The same rationale can be found in *Potter v. State*, 416 So. 2d 773, 777 (Ala. Crim. App. 1982) (leaving to jury's discretion weight to be given toolmark expert's opinion regarding ax used in murder).

217. 456 A.2d 561 (Pa. Super. 1983).

218. *See id.* at 565.

219. *See id.* at 566.

220. *Id.*

victim to the exclusion of all others.<sup>221</sup>

*Ramirez v. State*<sup>222</sup> is a rare instance of a court rejecting the opinion testimony of tool mark experts on the ground that a scientific basis for the testimony had not been established. In this case, the government sought to prove that the defendant's knife was the murder weapon by showing that the knife matched microscopic striations on the victim's cartilage, to the exclusion of all other knives: "[W]e find that no scientific predicate was established from independent evidence to show that a specific knife can be identified from the marks made on cartilage."<sup>223</sup> Cases such as this, requiring an adequate scientific basis to be established before allowing opinions based on that scientific predicate to go to the jury, have been rare in the judicial review of tool mark expert evidence.

In all, judicial opinions purporting to examine the scientific evidence on which tool mark expertise rests are remarkably few and the resulting opinions tellingly uninformative. While it is apparent that expert evidence on tool marks and firearms identification is universally admissible today, it is equally obvious that this universal admissibility has been accomplished without judicial evaluation of the validity of the underlying science or its application.

The absence of intelligible scrutiny is made unusually clear by the *Fasick* and *Clark* opinions. In the space of six months, in a pair of cases remarkably similar on their facts, a state supreme court went from confidently excluding to confidently admitting tool mark expert testimony, and in neither opinion did the court explain what led to its conclusions.<sup>224</sup> These opinions suggest that the judges do not lack strong feelings about the evidence, but they do lack sufficient understanding of the asserted basis of the opinion to enable them to explain why they are persuaded of its validity. These are disbelievers, or believers, without understanding.

Appellate courts have a duty not only to make decisions but to give the best explanation they can of their reasons for the decision. The absence of explanations undermines the opinion (no reasoning

---

221. MOENSSENS ET AL., *supra* note 4, at 379-80. See also James Starrs, *Procedure in Identifying Fingernail Imprint in Human Skin Survives Appellate Review*, 6 AM. J. FORENSIC MED. & PATHOLOGY 171, 172-73 (1985) (explaining that the presence of class, but not individual, characteristics, meant that a large number of fingernails in addition to the defendant's would have matched equally well).

222. 542 So. 2d 352 (Fla. 1989), *rev'd*, 651 So. 2d 1164 (Fla. 1995).

223. *Id.* at 354-55. The court specifically rejected *State v. Churchill*, 646 P.2d 1049 (Kan. 1982), because the Kansas court admitted toolmark identification evidence without first establishing the necessary scientific reliability predicate. See *Ramirez*, 542 So. 2d at 355. The *Ramirez* exclusion, however, was reversed on appeal. See *Ramirez v. State*, 651 So. 2d 1164 (Fla. 1995).

224. See *supra* notes 192-212 and accompanying text.

cannot be good reasoning) and deprives other courts of the incremental benefits of the thinking of prior courts that have struggled with the same problem. Being explicit about understanding and reasoning allows commentators and other courts to evaluate whether the opinion is correct or incorrect. To expose the wisdom or foolishness of one's thinking to the scrutiny of others may fill one with trepidation, especially when venturing into unfamiliar territory, but it is necessary to do if the law is to remain public and to work its way toward being more rational.

Silence about reasons is unusual behavior for appellate courts, which rarely are at a loss for explanations of their reasoning. This seems to further suggest that the courts recognized their duty to decide but either lacked the capacity to reason about the evidence or were not adequately informed by the parties. Courts have tools they can employ to try to better inform themselves: a visit to the library to examine the scientific literature, special masters, and requiring the parties to re-brief the court, among others. However great the challenge of such materials, the court's duty is to find a way to meet the challenge, not avoid it.

Later appellate courts evaded the responsibility of gatekeeping altogether, deferring either to the discretion of the trial court or to the jury, or by suggesting that there are no limits to the admissibility of expert testimony. As discussed earlier,<sup>225</sup> the decision to categorically exclude or admit scientific evidence is a decision of law, of legislative fact not adjudicative fact, and it therefore falls squarely within the province of the reviewing court. Moreover, the trial court has an independent responsibility to evaluate the admissibility of expert testimony. The jury's task, and its responsibility to try to understand and weigh the evidence, will come soon enough. But because it is coming, the judge is not excused from having to perform the court's gatekeeping role. Arguments that because there is expertise on everything, everything must therefore come in, are not likely to be accepted by any court any more. Lines must be drawn and articulated by courts or legislatures. To flesh out the contours of those lines, courts have a duty to explain how they determined on which side of that line any given asserted expertise fell.

Finally, at least one of these cases illustrates an attempt to use a legal tool to resolve scientific uncertainties when the tool simply is not up to the job. The *Fletcher* court's conflation of an attack on the scientific foundation of an expert's testimony with an attack on the expert's credibility<sup>226</sup> fails to recognize that credibility is incapable of assessing

---

225. See *supra* notes 47-53 and accompanying text.

226. If the *Fletcher* court were correct, then the entire body of law and commentary on the issue of the admissibility of scientific evidence, including Federal Rule of Evidence 702, could simply be dropped into the wastebasket and the whole matter referred to the jury.

the validity of science. It is one thing to evaluate the truthfulness of lay witnesses on issues of adjudicative fact by resorting to their credibility. But the existence or meaning of a body of asserted scientific knowledge does not depend on the credibility of the witness. Such a stance betrays not only ignorance of the scientific subject matter but also hopelessness about ever being able to understand it. Contrast this to a court hearing the legal arguments of counsel. The court does not ask itself if counsel is "credible," it asks itself if the arguments are sound. Science is like that to scientists. In listening to colleagues debate an issue, scientists do not ask themselves which side is more credible, they ask themselves which side is making the most persuasive arguments, given the body of empirical evidence and theory that are the materials they have to work with. Courts should be evaluating science much as scientists do, and much as they evaluate the arguments of lawyers. For this job, credibility is a useless tool.<sup>227</sup>

#### **D. Voiceprints: Judicial Cacophony**

The claim of "voiceprint" identification is that each person's vocal apparatus is unique and therefore the voice sounds each person makes are unique. By converting voice sounds to a visual display, using spectrographic equipment, the examiner can compare tracings and determine the identity of the person who uttered the questioned voice.

Judicial opinions on the admissibility of such voice identification expertise are widely divided. At present, by my count, expert testimony based on voice spectrographic analysis is admissible in six states<sup>228</sup> and excluded in eight,<sup>229</sup> admissible in four federal circuits<sup>230</sup> and excluded in one.<sup>231</sup> No consistent or coherent judicial view has evolved over time nor does a consistent view appear likely to emerge in the foreseeable future. The history of the courts' divergent responses to voice spectrography is instructive.

---

227. The one exception to this would seem to be with respect to the credibility of a study (not the witness). One may not believe the findings of a particular study. For example, one may think that a study funded by a particular interest biases the study. If so, that presumably is because the study was designed to produce favorable results, rather than a fair test, and a researcher should be able to see that in the design, so it still is a matter of intelligent scrutiny, not mere credibility. The worst situation would be where one does not believe a study because one thinks the researcher lied about the findings, and reported false results. That is presumed to be rare because nothing could be more destructive to the entire institution of science than "fudging," or "dry-labbing." The fact that there are words for such intellectual crimes suggests that they do occur.

228. Florida, Maine, Massachusetts, Minnesota, Ohio, Rhode Island. See Appendix.

229. Arizona, California, Colorado, Indiana, Maryland, Michigan, New Jersey, and Pennsylvania. See Appendix.

230. The Second, Fourth, Sixth, and Seventh. See Appendix.

231. The District of Columbia. See Appendix.

The accompanying table of Scientific Voice Identification Cases (see Appendix) lists the major voice identification opinions in chronological order, along with other information about the cases.

First, the Table reveals that there is no greater agreement in recent years than there was in the earliest days of voice spectrography. Of the first ten courts to consider the technique, six admitted it and four excluded it. The most recent ten to consider it were evenly divided on admission or exclusion. Second, we can see that the legal test of admissibility applied by the courts is highly correlated with the holding.<sup>232</sup>

Of the courts that claimed to apply the *Frye* test in a broad fashion—that is, treated the relevant scientific community as consisting of a range of relevant fields<sup>233</sup> and not merely the one narrowly concerned with performing the particular application that constitutes the technique at issue—only one admitted expert testimony of voice identification.<sup>234</sup> And vice-versa. Of courts that employed the *Frye* test narrowly—narrowing the relevant scientific field to “those expected to know,” that is, those which perform the specific application at issue—not one excluded the testimony.<sup>235</sup> These two versions of the *Frye* test, and their predictably opposite conclusions, illustrate one of the familiar criticisms of *Frye*, namely, that defining the relevant scientific fields broadly or narrowly largely dictates the conclusion that will be reached.

Of courts that claimed to employ a “relevancy” or “reliability” test—frequently equated, at least in the past, with the test embodied in the Federal Rules of Evidence—thirteen admitted<sup>236</sup> voice identification expert testimony and three excluded<sup>237</sup> it. What courts required for the expertise to be sufficiently “reliable” varied considerably. Most concluded that as long as there was something to be said on behalf of voice identification, that was enough to let it in. One court noted only that the witness was a credentialed expert and cited other jurisdictions that had admitted such testimony.<sup>238</sup> Using a similarly minimal threshold, however, another court excluded the evidence, concluding that its almost presumptive reliability was outweighed by its risk of being given

---

232. Notice that I merely say “correlated.” I venture no guess as to which came first: the rule to be applied, or the conclusion about admissibility.

233. Concerning voice spectrography, that could mean acoustical engineering, anatomy, electrical engineering, linguistics, phonetics, physics, physiology, psychology, and statistics because the technique of voice spectrography made assumptions about or borrowed principles from each of these fields.

234. See cases cited in Appendix.

235. See cases cited in Appendix.

236. See cases cited in Appendix.

237. See cases cited in Appendix.

238. See *United States v. Smith*, 869 F.2d 348, 352-53 (7th Cir. 1989).

excessive weight by factfinders.<sup>239</sup> Yet another court gave the scientific evidence on the proffered expertise a close and thoughtful examination, much like what the *Daubert* gloss on the Federal Rules would seem to require, and concluded that voice identification expert testimony was inadmissible.<sup>240</sup>

Finally, there were courts that employed the test suggested by McCormick, that is, admissibility based essentially on logical relevancy under Federal Rule of Evidence 401, and exclusion if probative value was substantially outweighed by the concerns invoked by Federal Rule of Evidence 403.<sup>241</sup> Every court employing this kind of test found voice identification expert testimony admissible.<sup>242</sup>

Some opinions reached their conclusions without employing a discernible legal test.<sup>243</sup>

The mere fact that some courts refused to admit voice identification expert evidence is itself significant in light of the traditional receptiveness of the courts to forensic individuation techniques. Why has voice identification been treated differently? Several interconnected answers are plausible.

One may be that more judges have grown more thoughtful and discerning and less credulous about scientific offerings than their judicial ancestors had been. Numerous courts evaluating asserted voice identification expertise were critical of witnesses (testifying in favor of admissibility) who were mere technicians rather than educated scientists,<sup>244</sup> or whose livelihoods depended upon continued admission of the technique,<sup>245</sup> or who came from a very small and incestuous circle of proponents of the technique.<sup>246</sup>

---

239. See *State v. Free*, 493 So.2d 781, 787-88 (La. App. 1986).

240. See *People v. Collins*, 405 N.Y.S.2d 365 (Sup. Ct. 1978).

241. McCormick's test for scientific evidence admissibility states that:

General scientific acceptance is a proper condition for taking judicial notice of scientific facts, but it is not a suitable criterion for the admissibility of scientific evidence. Any relevant conclusions supported by a qualified expert witness should be received unless there are distinct reasons for exclusions. These reasons are the familiar ones of prejudicing or misleading the jury or consuming undue amounts of time.

MCCORMICK, HANDBOOK OF THE LAW OF EVIDENCE § 203, at 608 (3d ed. 1984).

242. See cases cited in Appendix.

243. See cases cited in Appendix.

244. "[The expert witness's] qualifications are those of a technician and law enforcement officer, not a scientist." *People v. Kelly*, 549 P.2d 1240, 1250 (Cal. 1976) (emphasis in original).

245. See *id.* at 1249. Compare this to the narrow version of the *Frye* test, which essentially asks the practitioners of a technique if they have sufficient confidence in their own work that they should be allowed to continue earning income from it.

246. See *id.* Of course, these shortcomings do not distinguish talker identification from most other forensic individuation techniques when they were gaining admission to the courts. Indeed, all but the third criticism continues to be true for them.

Another factor is that the literature of voice spectrography, both supporting and questioning the technique, was more quantified and qualified<sup>247</sup> than earlier courts had received about earlier forensic individuation techniques. This is because many of the people involved in voice identification came from fields that had a tradition of empirical testing of their ideas.<sup>248</sup> Indeed, more research was available to the courts about voice identification expertise than for any forensic individuation field that preceded it.<sup>249</sup> This immediately provided the courts with unusual resources with which to comprehend the imperfections of the technique.<sup>250</sup> In providing rigorous self-criticisms of its own concepts and methods (which all good science does), a field aids the courts greatly in making a more informed and sober assessment of the field and its likely contribution to the judicial fact-finding process.<sup>251</sup> Moreover, a field that develops out of an industrial or academic research tradition tends to conduct more research; a greater volume of research tends to produce a more complex and qualified impression of the technique in the mind of the court.<sup>252</sup>

In the face of actual data, the courts had a real choice to make. Although the technique could reduce uncertainty in identification, it also was demonstrably less than perfect. Errors were going to be made, and, unlike other fields of forensic individuation, voice identification proponents said so.<sup>253</sup> The courts had concrete error rates to

---

247. "Qualified" in the sense of limited, restricted, or circumspect. See Raymond Kent & Michael Chial, *Talker Identification*, § 25-2.0, in *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* (David L. Faigman et al. eds., 1997) (reviewing the empirical research).

248. For evidence of this, the reader is referred to the nearest university library to examine a journal article in acoustical engineering, linguistics, perceptual psychology, etc., and to compare them to articles in the forensic sciences.

249. Compare the various forensic identification sciences reviewed in FAIGMAN ET AL., *supra* note 15.

250. The same was true for DNA typing, and was not true for most other forensic individuation techniques.

251. When other fields lack such critiques, is that because there is nothing to question? Or because an uninformed and unquestioning consensus developed among members of the field? And how can courts distinguish between the two possibilities?

252. This presents a paradox: All else equal, the better a field studies and critiques itself, the more skeptical the courts' impression of it will be. The less a field understands what it is doing and what the limits are on its own capabilities, the more positive an impression the courts will develop of the field. For a number of the more conventional forensic individuation techniques, there is still no tradition of self-scrutiny or a literature of the results of rigorous testing which can inform the courts. At least in terms of their continued acceptance by the courts, those fields have nothing to gain and much to lose by adopting a tradition of inquiry, testing, and skepticism.

253. "Possibly, no combination of methods may ever produce absolutely positive identifications or eliminations in 100% of the cases submitted." Oscar I. Tosi, *The Problem of Speaker Identification and Elimination*, in *MEASUREMENT PROCEDURES IN*



evaluate.<sup>254</sup> How good is good enough? How much error is too much? The law provides no standards for making that assessment. A technique that committed five false positive errors for every 100 decisions may be viewed by some courts as more than adequate and by other courts as not nearly good enough.

Finally, the courts may have been overwhelmed by the studies. Although more research means a greater potential to understand the scientific questions at issue, it also may have confused some (or many) courts, which have limited capacity and limited time to interpret and evaluate the empirical studies. If this were the problem, help was on the way.

Unique<sup>255</sup> assistance in evaluating the available data came into being only a decade after voice identification made its first appearance in the courts. Help came in the form of a careful review of voice spectrography by the National Academy of Sciences. A panel of highly knowledgeable scientists from diverse relevant fields was formed; they carefully reviewed the relevant scientific literature; and they published their report.<sup>256</sup> Although the Report wisely declined to make a specific recommendation concerning admissibility—appreciating that to do so called for value judgments<sup>257</sup> that should be made by the law, not usurped by scientists—the Report nevertheless could not be said to have given a good report card on the state of voiceprint analysis. Among other things, it concluded:

[The assumption] that intraspeaker variability is less than . . . interspeaker variability . . . is not adequately supported by scientific . . . data.<sup>258</sup>

\* \* \*

Estimates of error rates now available pertain to only a few of the many combinations of situations encountered in real-life situations. These estimates do not constitute a generally adequate basis for a ju-

---

SPEECH, HEARING, AND LANGUAGE 399, 428 (Sadanand Singh ed., 1975).

254. See *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 594 (1993) (stating that an admissibility criterion for a technique is its "known or potential rate of error."). Until very recently, such information was completely unavailable for most forensic science. See PETERSON ET AL., *supra* note 96.

255. Up until that time. There have subsequently been two NAS panels formed to review the data on the technique of DNA typing. See NATIONAL RESEARCH COUNCIL COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, DNA TECHNOLOGY IN FORENSIC SCIENCE (1992); NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996).

256. See COMMITTEE ON EVALUATION OF SOUND SPECTROGRAMS, NATIONAL ACADEMY OF SCIENCES, ON THE THEORY AND PRACTICE OF VOICE IDENTIFICATION (1979).

257. Most notably: How much uncertainty or known rates of error are disqualifying for the purposes of trial?

258. COMMITTEE ON EVALUATION OF SOUND SPECTROGRAMS, *supra* note 256, at 2.

dicial or legislative body to use in making judgments concerning the reliability and acceptability of aural-visual voice identification in forensic applications.<sup>259</sup>

Upon publication of the report, the FBI ceased performing voice identification for the purpose of offering testimony in court,<sup>260</sup> and it was expected<sup>261</sup> that the courts would stop admitting voice identification expert testimony, at least until the scientific support for it improved.<sup>262</sup> Of the eleven judicial opinions written since release of the NAS report, however, half admitted the expert testimony.<sup>263</sup> Still more troubling, most of those opinions make no mention whatever of the NAS review,<sup>264</sup> and only one of the two that did cite it appears actually to have read it.<sup>265</sup> In short, for the most part, the courts decided the post-NAS cases just as they decided the pre-NAS cases, that is, as if the report did not exist.<sup>266</sup>

Only among the voice identification cases do we see modern courts rejecting claims of forensic identification expertise in large numbers. In attempting to account for this, it is tempting to suppose that judges have become more skeptical or better informed, or that the lawyers practicing before them have. Perhaps on appreciating the greater complexity of a claimed identification science, or better understanding how difficult it is to prove the empirical claims, or on recognizing the difference between technicians and scientists, or the possible strains on the objectivity of members of fields whose very existence depends on the credulity of judges, many courts are not persuaded.

Several considerations make this explanation doubtful, at least as a general matter.<sup>267</sup> One is the opinions themselves. Reading them does not suggest to a reader much greater scientific literacy than existed earlier in the century. A second reason is the extraordinarily different fate of the next field to be considered, bite mark identification,

---

259. *See id.* at 60.

260. *See* MOENSSENS ET AL., *supra* note 4, at 645. But, as with the polygraph, they continued doing spectrographic tests for investigative purposes. *See id.* at 645 n.30.

261. *See id.* at 645.

262. Few if any of the scientific shortcomings raised by the report have been solved by subsequent research. *See* Kent & Chial, *supra* note 247, § 25-2.3.

263. *See* cases cited in Appendix.

264. *See* cases cited in Appendix.

265. *See* State v. Gortarez, 686 P.2d 1224, 1235-36 (Ariz. 1984). The *Maivia* court's discounting of the NAS report based on the idea that it had grown dated assumed that more informative research had changed the picture. Even to the present, however, that has not come about. *See* Kent & Chial, *supra* note 247, § 25-2.3.

266. Whether this reflects the shortcomings of counsel (for not drawing the courts' attention to the NAS study and the significance of its findings) or the courts (for not finding it themselves, or not appreciating its value to their decisions), I am unable to say.

267. That is to say, some individual opinions are impressive in their thoughtful grasp of the scientific issues. *See, e.g.,* People v. Collins, 405 N.Y.2d 365 (Sup. Ct. 1978).

which made its entry into the courts at about the same time.<sup>268</sup> And, perhaps most convincing, is the massive failure of the courts after 1979, or the lawyers practicing before those courts,<sup>269</sup> to take note of, and then either learn from or thoughtfully reject, the National Academy of Sciences study on the very subject these courts were called upon to rule.<sup>270</sup>

What actually may distinguish the voice identification cases is that, unlike all of the other forensic identification sciences<sup>271</sup> the courts had seen, the founders of spectrographic voice identification produced studies which could be presented to the courts, showing weaknesses as well as strengths of the technique. The founders made the "mistake" of conducting studies on their subject and on their own performance because, in contrast to the other forensic identification sciences, they came from fields accustomed to conducting empirical research and basing their conclusions on hard data: acoustical engineering, electrical engineering, linguistics, physiology, experimental psychology. Data always are messier than theoretical suppositions, and will always give more pause. Ironically, then, it may be that any field that brings more science with it will be viewed more skeptically. The other forensic identification sciences may have done nothing better to help their cause, at least in the eyes of courts, than to do little or no empirical research.

Finally, the voiceprint cases also illustrate some things about the relationship of admissibility doctrine and admissibility decisions. First, they illustrate the malleability of the *Frye* test, by which judges can cast their net as narrowly or widely as they choose in order to find the "particular field in which it belongs" and within which the technique at issue must have been generally accepted and thereby admissible. As we can see with these voiceprint cases, those courts that cast a wide net almost always excluded the asserted expertise and those casting a narrow net admitted it. More generally, we see a correlation between the rule applied and the result reached. This can give us confidence in the

---

268. See *infra* subsection III.E. concerning bite marks.

269. I have not examined the briefs filed with the courts to see if the attorneys cited the NAS study.

270. What lesson is this? Lawyers and judges, apparently, do not research particularly well the scientific materials bearing on their cases. How might a patient feel about being operated upon by a surgeon who was unaware of the single most pivotal review of research concerning the operation being performed? How would any client, or the larger society, feel about their legal decisions being made by judges and lawyers who were unaware of the single most pivotal study concerning the case being decided? Perhaps the courts that cited but did not read (or read but did not comprehend) the NAS Report should be even more troubling.

271. With the exception of DNA typing, which came later, and which similarly seems to prove the "rule" suggested here.

rule of law or not, depending upon whether the courts followed an established rule to its inevitable consequence, or chose a rule to bring about the result desired. Often, the applicable rule governing admissibility of asserted scientific evidence has been vague and unclear, so that courts have considerable wiggle room with which to select the rule to use in any given case.<sup>272</sup>

#### E. Bite Marks: Exempt from *Frye* and *Collins*

The analysis of bite marks for the purpose of identifying a criminal perpetrator is a specialized task within the broader discipline of forensic odontology. Forensic dentists traditionally had been called upon to identify the remains of disaster victims by comparing the victims' dentition with dental records.<sup>273</sup> In trying to identify perpetrators of crime, the forensic dentist seeks to compare a suspect's dentition with a latent mark left in the victim's flesh or in some edible substance found at the scene of a crime.<sup>274</sup>

Bite mark identification is one of the newest areas of forensic identification. The courts have been virtually unanimous in admitting such testimony. Bite mark expert opinion has now been admitted in most jurisdictions of the United States, the great majority of those occurring since 1980.<sup>275</sup>

Several ironies accompany this legal history. One is that forensic odontologists, perhaps reflecting a grounding in scientific skepticism that is absent from the traditional forensic identification sciences,<sup>276</sup> were more doubtful about whether the state of their knowledge permitted them to successfully accomplish the challenging task of identifying a perpetrator "to the exclusion of all others." The history of other areas of forensic identification reveals few with similar self-doubts. Second, in spite of the existence of these profound doubts, the courts began admitting expert testimony on bite marks even under the

---

272. See Bohan & Heels, *supra* note 38, at 1036, Table 2 (comparing variety of "tests" adopted among and within states for determining admissibility of purported scientific evidence).

273. See MOENSSENS ET AL., *supra* note 4, Chapter 16.

274. The two tasks differ in important ways. In the disaster situation, there is a finite number of candidates to identify, and full dentition often or usually is available both from the victims and the dental charts. In bite mark cases, the number of potential suspects is huge, the bite marks include only a limited portion of the dentition, and flesh provides a far less clear medium for recording a bite mark than having the teeth (of the disaster victim) themselves.

275. For a compilation of these cases, see DAVID L. FAIGMAN ET AL., *SUPRA* NOTE 15, § 24-1.0.

276. By traditional forensic sciences, I mean fingerprints, footprints, toolmarks, handwriting, firearms, and so on.

*Frye* test.

And, third, rather than the field convincing the courts of the sufficiency of its knowledge and skills, admission by the courts apparently convinced the forensic odontology community that, despite their doubts, they really were able to perform bite mark identifications.<sup>277</sup>

The first case in the United States to confront the admissibility of expert testimony on bite mark identification was in the Texas case of *Doyle v. State*.<sup>278</sup> Doyle was charged with burglary. At the site of the burglary was found a piece of partially eaten cheese. After arresting Doyle, the sheriff asked him to bite a piece of cheese, which the suspect voluntarily did. A firearms examiner then compared the two pieces of cheese to try to determine if the questioned and the known tooth marks had been made by the same person. The firearms examiner concluded that they had. At trial a dentist testified that from his own examination of plaster casts of the cheese bite marks, he also held the opinion that one and the same dentition had made both sets of bites.<sup>279</sup> The court upheld the admission of this bite mark opinion testimony.<sup>280</sup>

Although the empirical research necessary to form the scientific ground for such a conclusion had not yet been accomplished,<sup>281</sup> the defense in *Doyle* did not contest admissibility by raising any issue of scientific validity, but instead raised only procedural challenges.<sup>282</sup> Be-

---

277. In their book on scientific evidence, MOENSSENS ET AL., *supra* note 4, conclude the following about the relationship between the courts and expert opinion on bite mark identification:

The wholesale acceptance, by the courts, of testimony on bite mark identification has transformed the profession. Whereas prior to 1974 the main thrust of forensic dentistry was to prove identity of persons by means of a comparison of postmortem and antemortem dental records in mass disasters, the profession has changed direction and is now heavily involved in assisting prosecutors in homicides and sex offense cases. Having received judicial approval of bite mark comparisons, there seems to be no more limit on the extent of forensic odontological conclusions.

*Id.* at 985.

278. 263 S.W.2d 779 (Tex. Crim. App. 1954). Although this was the first appellate consideration of bite mark evidence, the technique had been used for related identification purposes for decades. See MOENSSENS ET AL., *supra* note 4, § 16.04.

279. See *Doyle*, 263 S.W.2d at 779.

280. See *id.* at 780. The court did not consider the admissibility of the bite mark on evidentiary grounds, but on self-incrimination grounds. See *id.* at 779.

281. See Raymond Rawson, *Identification from Bite Marks*, § 24-2.1.1[1], in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (David L. Faigman et al. eds., 1997) (The earliest relevant empirical research was conducted two decades after *Doyle* was decided.).

282. See *Doyle*, 263 S.W.2d at 779-80. Instead, the defense raised only the issue of whether obtaining the bitten cheese from the defendant constituted a confession and thereby violated a Texas statute prohibiting obtaining confessions without warning defendants of

cause the issue never arose, the *Doyle* opinion contained no holding on the fundamental question of the scientific status of bite mark identification. Nevertheless, another Texas court relied on *Doyle* twenty years later as the basis for rejecting an appellant's contention that bite mark test results were of unproven reliability.<sup>283</sup>

The cornerstone case on the admissibility of bite mark identification was decided in 1975, in California. In *People v. Marx*,<sup>284</sup> the court undertook to grapple with the real issues of admissibility raised by bite mark identification, but succeeded only in avoiding them. *Marx* involved the brutal murder of an elderly woman.<sup>285</sup> An elliptical laceration was discovered on the victim's nose.<sup>286</sup> This mark was judged to be a human bite mark. Impressions were made of the wound for comparison with a cast of the defendant's teeth.<sup>287</sup>

At trial, three odontologists testified that in their opinion the defendant's dentition matched the bite wound.<sup>288</sup> One of those experts took pains to note that in many other cases he had refused to offer a firm opinion or even to testify about an identification. This case, however, was an exception in that the dentition at issue was unusual and the bite mark was "exceptionally well defined." The witness characterized these bite impressions as the clearest he had ever seen, either personally or in published literature.<sup>289</sup> Despite the expert's caution, and unusual case facts emphasizing the rarity of both the dentition and the bite marks, once the courtroom door was pried open, *Marx* became the admission ticket for a far wider and more dubious array of dentition evidence in many subsequent cases.

On appeal, the defense challenged the admission of expert opinions on bite wound identification on the ground that the purported skills were not sufficiently established or generally accepted in the field of forensic dentistry.<sup>290</sup> California law presented two major barriers to admission. First, California followed the *Frye* test, so that the prosecution had to show that the field of dentistry or perhaps just forensic dentistry, generally accepted the theory and techniques of bite mark iden-

---

their likely use.

283. See *Patterson v. State*, 509 S.W.2d 857, 863 (Tex. Crim. App. 1974). If the admission of the bite mark was in fact error, the court stated that it was harmless error given the appellant's written confession that he had indeed bitten the victim.

284. 54 Cal. App. 3d 100 (Cal. Ct. App. 1975).

285. See *id.* at 351.

286. See *id.*

287. See *id.* at 352.

288. See *id.* at 353 n.5. For a detailed discussion of the scientific evidence in this case, see Gerald L. Vale et al., *Unusual Three-Dimensional Bite Mark Evidence in a Homicide Case*, 21 J. FORENSIC SCI. 642 (1976).

289. See *Marx*, 54 Cal. App. 3d at 108 n.8 (statement of Dr. Vale).

290. See *id.* at 353.

tification.<sup>291</sup> Second, the California Supreme Court had prohibited speculative probability estimates from forming the basis of a criminal conviction.<sup>292</sup>

The appeals court acknowledged that there was "no established science of identifying persons from bite marks . . . ." <sup>293</sup> Moreover, the theory of bite mark identification, like all forensic identification science, is based on estimations of the probability that two or more people could leave the same bite mark.<sup>294</sup> If that probability is sufficiently low, the forensic dentist will testify that the defendant's teeth are the ones that left the crime scene marks. Yet no data existed on those probabilities. The state of forensic odontology would, on the face of it, seem to have left it unable to scale these legal barriers. Nevertheless, the court of appeals reasoned its way over these hurdles and upheld admission. Here is how:

The Court of Appeals deflected the implications of *Frye*, by interpreting that test in these terms:

The *Frye* test finds its rational basis in the degree to which the trier of fact must accept, on faith, scientific hypotheses not capable of proof or disproof in court and not even generally accepted outside the courtroom. *Frye*, for example, involved the lie detector test in which the trier of fact is required to rely on the testimony of the polygrapher, verified at most by marks on a graph, to which the expert's hypothesis gives some relevant meaning. [Other cases reflect] [t]he same concern that the trier of fact will be overwhelmed by 'ill conceived techniques with which the trier of fact is not technically equipped to cope,'<sup>295</sup> sacrificing its independence in favor of deference to the expert.

The court thus distinguished *Marx* from *Frye* by reasoning that *Frye* applied only to evidence that was indecipherable without an expert's interpretation, whereas *Marx* involved models, photographs, X-rays, and slides of the victim's wounds and the accused's dentition, all of which were clearly visible for the jury to view, assess, and verify on their own during court proceedings, without having to rely on the expert odontologist as a necessary intermediary.<sup>296</sup> Forensic odontologists, no doubt, would be surprised to learn that once the pictures are taken and the models cast, no special expertise or judgment is required to assess whether or not the wound was made by the defendant's dentition or by someone else's.

---

291. See *id.* at 355-56.

292. See *id.* at 357 (citing *People v. Collins*, 438 P.2d 33, 41 (Cal. 1968)).

293. *Id.* at 353.

294. See Rawson, *supra* note 281, at § 24-2.1.1[1].

295. See *Marx*, 54 Cal. App. 3d at 111 (quoting *People v. Collins*, 438 P.2d. 331, 41 (Cal. 1968)).

296. See *id.* at 356.

Alternatively, the *Marx* court concluded that *Frye's* requirements *had* been met because the methods used to facilitate the bite mark identification were not really novel:

[T]he experts did not rely on untested methods, unproven hypotheses, intuition or revelation. Rather, they applied scientifically and professionally established techniques—X-rays, models, microscopy, photography—to the solution of a particular problem which, though novel, was well within the capability of those techniques.<sup>297</sup>

On this view, *Frye* is about tools, not the meaning of the information collected with the help of the tools.<sup>298</sup> By this analysis, even the polygraph at issue in *Frye* in 1923 would have passed the *Frye* test. While the reliability of the tools is by no means unimportant, that is not the central question in bite mark identification. The most fundamental issues in bite mark identification, as with all forensic identification, are (a) whether the population variation in the relevant characteristics is immense (ideally, infinite, so that each dentition truly is unique), (b) whether in practice that underlying variation is adequately captured by the available tools and evidence, and (c) whether the latent mark has enough distinct variation in it to allow the probability that other dentition may have left the mark to fall comfortably low.<sup>299</sup> None of this essential knowledge was, or usually is, available to the jury. The question the court should be focusing on, however, is whether it is available to the expert.<sup>300</sup>

Finally, the *Marx* court appears to have believed that evidence on or even thoughts by the expert about the probabilities underlying bite mark identifications would be inadmissible in California under the rule of *People v. Collins*.<sup>301</sup> It held that, since the experts never had actual data and did no calculations of probability, but instead remained impressionistic and intuitive, “[n]one of the witnesses engaged in a ‘trial by mathematics’ [citing *Collins*] on or off the witness stand. There was no error.”<sup>302</sup>

---

297. *Id.* at 356 (footnotes omitted).

298. On this theory, of course, the polygraph in *Frye* would have been admissible, since the physiological measures “applied [were] scientifically and professionally established techniques.” *Id.* at 356. The central scientific issue in *Frye* was what one could infer from those measures. See *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir 1923).

299. See Rawson, *supra* note 281, nn. 14-17 and accompanying text.

300. If it is not, in some form or fashion, then the expert is speculating.

301. See *Marx*, 54 Cal. App. 3d at 112, 126 Cal. Rptr. at 357 (citing *People v. Collins*, 438 P.2d 33).

302. See *id.* (quoting *Collins*, 438 P.2d at 41). Recall that earlier in the opinion, when deflecting the *Frye* challenge, the court assured itself that “intuition or revelation” played no part in bite mark identification. *Id.* at 356. Now, in deflecting a challenge based on *Collins*, the court thought it had to assure itself of the opposite, that experts dealt with the probabilistic underpinnings of bite mark opinions only through intuition and without any explicit statistical calculations.



First of all, the *Marx* court overstates the *Collins* prohibition. In *Collins*, a statistically based identification was inadmissible because it was based on (a) speculative probabilities (rather than known relative frequencies of the attributes at issue),<sup>303</sup> (b) a lack of proof that the attributes of interest consisted of independent events,<sup>304</sup> and (c) invalid conclusions of rarity inferred from the probability calculation, which were suggested to the jury as establishing proof beyond a reasonable doubt.<sup>305</sup> Absent these flaws, such evidence presumably would be admissible. Second, whether a jury will be allowed to hear the numbers is one thing. Whether they may be heard *by a court* in deciding a preliminary question such as the admissibility of purported scientific evidence is quite another. Any court following Federal Rule of Evidence 104 or an equivalent rule<sup>306</sup> plainly is authorized to do so,<sup>307</sup> and any court following *Daubert* apparently has a positive *duty* to do so.<sup>308</sup>

Most important, perhaps, *Marx* is one of the rare cases to realize that much forensic identification evidence invokes the same reasoning that the California Supreme Court found so troubling in *Collins*. To believe that the experts, or the court in deciding a question of admissibility of expert evidence, should eschew consideration of the underlying data and the statistical inferences to be drawn from those data, seems to stand *Collins* on its head. It does not solve any problem that *Collins* was concerned with; it merely pretends the problems are not there. The court is reading *Collins* as saying that the proper way for experts and courts to think about the statistical aspects of a technique is not to think about them, except in vague and cloudy terms. But the existence and nature of probability data are at the heart of the theory of forensic identification. Hiding one's eyes from precisely what needs the closest examination is no solution.<sup>309</sup>

---

303. See *Collins*, 438 P.2d at 38.

304. See *id.* at 39. In order to apply the multiplication rule to calculate the probability, each component should be uncorrelated with the other components.

305. See *id.* at 40.

306. See, e.g., CAL. EVID. CODE §§ 310, 400-403 (West 1995).

307. See FED. R. EVID. 104. Moreover, in deciding such preliminary questions, the court "is not bound by the rules of evidence except those with respect to privileges." FED. R. EVID. 104(a).

308. Among other data to be considered is a technique's "known or potential error rate." *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 594 (1993).

309. DNA evidence is the best example of a field attempting to provide the necessary data and probability calculations, and the chief exception to the pattern that evidence offered by the forensic identification sciences resembles the evidence offered in *Collins*. The *Marx* interpretation of *Collins* as prohibiting any explicit calculations of probability by experts would seem to exclude DNA typing because DNA experts collect, analyze, and explicitly rely on probability data. Notwithstanding the *Marx* court's reasoning, forensic odontologists appreciate that these sorts of data are necessary, and are developing them. See Rawson, *supra* note 281, § 24-2.1.1[1].

The following year, in 1976, Illinois had its first occasion to consider the admissibility of bite mark evidence. In *People v. Milone*, the Court of Appeals held it admissible as "a logical extension of the accepted principle that each person's dentition is unique."<sup>310</sup> The court based this on its earlier practice of using dental records and dentition to identify deceased persons.<sup>311</sup> In *Milone* expert witnesses disagreed sharply on the question of the validity and utility of bite mark identifications. The testimony of three forensic dentists was offered by the prosecution and four by the defense.<sup>312</sup> The defense experts testified and cited odontological literature showing, at the least, considerable disagreement among forensic odontologists as to whether offenders could be uniquely identified from bites left in the flesh of victims.<sup>313</sup> Notwithstanding the controversy, in the trial record and in the literature, the court still managed to find that the general acceptance standard had been met.<sup>314</sup> As virtually always happened, and contrary to the approach of *Daubert*, the *Milone* court held that questions about the truth of the proposition quoted above—of infinite variation and unique identifiability—went to the weight of the expert testimony, not to its admissibility.<sup>315</sup>

By 1978, a California Court of Appeals flatly held that the testimony of three forensic odontologists established that bite mark identification had gained the required general acceptance in the relevant scientific community.<sup>316</sup>

After the 1970's, there was no longer any debate in the courts—if ever there had been—about the admissibility of bite mark expert opinions to identify perpetrators of crime. Every subsequent appellate case concerning the issue has held bite mark identification testimony to be admissible. The rationale for admission in all of these cases has followed one of two general approaches. Some courts have followed the *Marx* precedent in holding that bite mark identification is generally accepted within the relevant scientific community.<sup>317</sup> Others simply

---

310. 356 N.E.2d 1350, 1358 (Ill. App. Ct. 1976). This is another illustration of Quetelet's hypothesis becoming accepted as a law of nature by the courts. See *supra* notes 69-71 and accompanying text.

311. See *Milone*, 356 N.E.2d at 1358 (citing *People v. Maddox*, 237 N.E.2d 845 (Ill. App. Ct. 1968)).

312. See *Milone*, 356 N.E.2d at 1355-56.

313. See *id.*

314. See *id.* at 1360. Incidentally, as a testament to the power of weak or inapt precedents, the court cited the Texas cases of *Doyle* (which had no data) and *Patterson* (which relied on *Doyle*), as well as California's *Marx* (which dealt with unusual dentition, in contrast to the apparently more common dentition of the present case). See *id.* at 1359.

315. See *id.* at 1360.

316. See *People v. Slone*, 76 Cal. App. 3d 611, 625 (Cal. Ct. App. 1978).

317. See, e.g., *Spence v. State*, 795 S.W.2d 743, 752 (Tex. Crim. App. 1990) (per curiam).

held that whether to admit such expert opinion lies within the discretion of the trial court.<sup>318</sup>

Perhaps the most unusual legal development in these cases has been a recent appeal by the defendant convicted in *People v. Milone*.<sup>319</sup> Paroled after serving nearly twenty years in prison for murder, Milone continues to insist upon his innocence and is seeking legal redress to clear his name. In federal court, under both the *Frye* and *Daubert* standards, he has challenged the original decision to admit the expert bite mark testimony.<sup>320</sup> Another murder victim was later found in the same area where the victim in the Milone case had been found. A potential bite mark from the second murder victim was linked to one Macek. The marks in the two cases were judged by at least one forensic odontologist to be indistinguishable from each other.<sup>321</sup> Macek signed a confession to having killed the victim in the *Milone* case, a confession that later was withdrawn.<sup>322</sup> For present purposes, more important than the question of whether or not Macek killed both victims, is the suggestion that the relevant portions of dentition of two different suspects were indistinguishably alike.<sup>323</sup>

Although the Court of Appeals for the 7th Circuit expressed sympathy with the accused's request, in light of the new evidence presented, it declined to rule on the case for lack of a constitutional basis for granting relief, and because principles of federalism precluded a federal court from re-examining an issue of fact that is reserved to the states.<sup>324</sup>

The bite mark cases suggest that the quality of a field's science can have less impact on the courts' decisions about the field than the courts' decisions about the field have on the field's beliefs about itself.

---

318. See, e.g., *Kennedy v. State*, 640 P.2d 971, 978 (Okla. Crim. App. 1982). Therefore, the trial court's admissibility determination is subject only to deferential review by appellate courts. On this view, of course, forensic odontology can be admitted as an exemplary science in one trial courtroom one day and excluded as charlatany in the courtroom next door the following day.

319. See *supra* notes 310-315 and accompanying text.

320. See *Milone v. Camp*, 22 F.3d 693 (7th Cir. 1994).

321. The forensic odontologist, later President of the American Academy of Forensic Sciences, had been a defense expert in *Milone* and wrote about these cases in Lowell Levine, *Forensic Dentistry: Our Most Controversial Case*, in *LEGAL MEDICINE ANNUAL* 1978 73 (Cyril H. Wecht ed., 1979).

322. See discussion in *Missouri v. Sager*, 600 S.W.2d 541, 571-72 (1980).

323. See Levine, *supra* note 321, at 77. Such findings are not unique in the identification sciences—notably document examination and bertillonage. In the present case, however, it might be noted that the comparison was not made using standard methods or procedures because of Macek's having undergone extraction of his teeth. The comparisons of the bite marks to x-rays of pre-extracted teeth hold little similarity to standard comparison procedures of overlaying biting edges onto the injury patterns on the skin.

324. See *Milone*, 22 F.3d. at 705.

As a field, forensic dentists appreciated the magnitude of the challenge of unique identification better than most identification scientists before them, and the need for, but lack of, data supporting the opinions they would like to have been able to offer, and therefore came to court with admirable self-doubt. But pivotal opinions like *Marx* and *Milone* swept aside those doubts.<sup>325</sup> The courts in these pivotal opinions also did not find admissibility doctrine to present much of a barrier. Patent disagreement within the field about their capability to individualize did not prevent courts from finding general acceptance.

Furthermore, the California appeals court that decided *Marx* managed to find a knothole in the fence that *People v. Collins* erected against probabilistic evidence that lacked a real database and relied upon erroneous<sup>326</sup> and misleading interpretations of probability.<sup>327</sup> That court seemed to sense how serious a threat *Collins* posed, not only to bite marks but to most of forensic identification science.<sup>328</sup> Perhaps the *Marx* court is correct in applying *Collins* only to explicit probability calculation—that is, it is permissible if the expert intuitively tells the jury only the conclusion they lead to. On the other hand, if *Collins* is still good law in California, it remains to be explained how any forensic identification science, other than DNA typing, can be admitted into evidence in California courts. That is, unless and until forensic scientists develop the empirical data needed to replace probabilistic guesstimations with data.<sup>329</sup>

## Conclusion

A great deal of intelligence can be invested in ignorance when the need for illusion is deep.<sup>330</sup>

What lessons might be drawn about the law's regulation of pur-

---

325. If the courts think we are good enough, then "by golly" we must be. Although there are now many forensic dentists who will testify, there also are many who continue to see important weaknesses in the theoretical and empirical foundations of their central claims, and continue to do research to try to fill those gaps and improve the quality of the work they do. See Rawson, *supra* note 281. See also *MANUAL OF FORENSIC ODONTOLOGY* (C. Michael Bowers & Gary L. Bell eds., 1995).

326. Erroneous in that the statistical model did not take into account whether or not the features of interest were independent, which has a large impact on the probability arrived at, whether explicitly or intuitively.

327. Misleading in that the jury is not helped to appreciate how diagnostic the probabilistic conclusions are (or aren't).

328. DNA typing is the exception because, in its collection and analysis of explicit data, DNA typing adheres to *Collins's* requirements.

329. See Saks & Koehler, *supra* note 57, at 372.

330. Saul Bellow, *quoted in* PETER'S QUOTATIONS: IDEAS FOR OUR TIME 428 (Laurence J. Peter ed., 1977).

ported scientific evidence from our review of the major formative encounters of the courts with the forensic identification sciences? The most obvious lesson is that this area of law is riddled with contradiction, confusion and chaos. Here is an area of law typified by uninformed acceptance or rejection (usually the former) of empirical claims that courts rarely made any effort to try to understand; by judicial opinions that failed to explain the courts' reasoning about the asserted expertise; by evasion or manipulation of the governing legal tests, or the use of no legal test at all; and by abdication of the intellectual and institutional duty to decide, in deference to the "authority" of the very witnesses whose offerings were to be evaluated.

The most ironic of these contradictions is that such a performance would occur in relation to scientific evidence. Science should not be mysterious. Indeed, science evolved as a means of taking the mystery out of our understandings of the way the world works. The scientific method is essentially a set of logical procedures for testing the validity of empirical claims.<sup>331</sup> Science is a system for making explicit the assumptions underlying beliefs and making plain the evidence and inferences on which conclusions stand (or fall).<sup>332</sup> Real science is not an alternative mysticism.

How can this potentially most lucid of evidence have been treated as if it were merely so much magic—impenetrable, calling for acts of faith rather than reason? How can what might have been expected to be the most helpful and most manageable kind of evidence have given rise to some of the most vacuous and opaque judicial opinions? Let's see if we can draw some lessons from this experience that help make sense of it.

### A. The Foreground Can Hide the Background

The most prominent issue in discussions of the admissibility of scientific evidence has been the choice of a rule. A well conceived and well drafted rule of evidence or procedure, it is assumed, should enable courts to admit valid while excluding invalid scientific evidence. If so, it would follow that the difficulties and controversies over the admissibility of scientific evidence during the past century are simply an indi-

---

331. This is true at the level of facts as well as at the level of theories, because theories in science are theories about empirical phenomena. All questions that science aims to deal with are empirical, because the only questions science is peculiarly competent to deal with are empirical.

332. For an extended discussion of the scientific method for lawyers and judges, see generally Michael J. Saks, *Scientific Method: The Logic of Drawing Inferences from Empirical Evidence*, in MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY (David L. Faigman et al. eds., 1997).

cation that rule makers have not yet hit upon the right test.

This perspective places emphasis on the truth-seeking function of trials. Valid scientific knowledge ought to help resolve factual uncertainties. If the courts, exercising their gatekeeping responsibility, can separate the valid from the invalid science, then factfinders should gain very real help.

From this viewpoint, the shifts from the implicit marketplace test to *Frye* and from *Frye* to *Daubert* are significant ones because they are experiments with different ways of answering the question: Which asserted knowledge is valid? Under the marketplace test, judges asked whether consumers of an asserted expertise found it useful in important everyday affairs of life outside the courtroom. Under *Frye*, judges asked whether the community of experts believed themselves to be in possession of valid and useful knowledge. Under *Daubert*, judges ask whether the asserted scientific expertise can be seen to be valid, based on its underlying empirical support. By definition, the tests before *Daubert* did not have judges evaluating science, but evaluating instead what the consumers or producers of an asserted science thought of it. The shift to *Daubert* clearly is a shift toward evaluating science more in the manner that scientists evaluate purported science, and therefore represents an increased emphasis on the truth-finding value of trials.

But these rules may conceal more than they reveal about scientific evidence admissibility decisions. Many of the judicial opinions we examined seemed to use no rule at all. Others would pick and choose the test or its variant to be applied, or manipulate the rule they felt obligated to use, or claimed to be using.<sup>333</sup> In short, it seems apparent that in making admissibility decisions, courts were guided by considerations that went beyond any "admissibility test," and that could easily override any admissibility test.

What then, in addition to rule-guided tests of admissibility, might be driving the admissibility decisions?

The very promise of scientific evidence—that it can commandingly resolve vital and troubling factual uncertainties, such as the identity of the perpetrator of a crime—heightens two competing fears that judicial gatekeepers must harbor. One is the fear of permitting rubbish to infect a trial, mislead a jury, and lead to the wrongful conviction of an innocent person. Counterposing that is the fear of overlooking evidence that would forcefully reveal the guilt of a true perpetrator of a serious crime. Thus, the central dilemma of trials, namely, the countervailing fears of convicting the innocent or acquitting the guilty, are reduced to an evidentiary decision. The admissibility decision is infused with much of the anxiety that the scientific evidence proposes to

---

333. See also *infra* subsection C.

obviate. Moreover, the court can resolve this gatekeeping dilemma only if it surmounts new and more challenging intellectual hurdles.

In many or most of the pivotal cases involving forensic identification science, the courts responded to these fears and challenges mostly by avoiding the necessity of making a real decision about the asserted science. This was accomplished in a variety of ways, such as by resting the decision on collateral legal grounds,<sup>334</sup> or by passing the decision on to the jury,<sup>335</sup> or by trusting the experts to do the court's gatekeeping for it.<sup>336</sup>

These strategies usually, or necessarily, result in admission of the evidence. The trial goal served by such strategies has elsewhere been referred to as an "exorcism of ignorance."<sup>337</sup> That is, whether valid science or not, the scientific expert's testimony will help the fact finder escape from the quandary of deciding whether a defendant is indeed the perpetrator. Evidence advances that goal, more evidence is presumed to be better than less evidence, and if it points strongly in one direction, the fact finder can make a far less conflicted decision and feel better about it.

Scientific evidence admissibility decisions also are opportunities to win, or retain, public confidence in the decisions of courts. The marketplace test was best adapted to this purpose. If the public, in the marketplace, came to believe in the value of a given expertise, then a court could enlist that confidence for use in its own decisions by adding the same expertise to the ingredients of its trials. *Frye* was the second best method of doing this. But in an era of specialization, perhaps courts could do no better in strengthening public confidence in the input to and outcome of trials than by piggy-backing their own admissibility decisions onto the beliefs of respected professional communities.<sup>338</sup>

---

334. For example, where a sample is sought from a defendant, by focusing on the issue of self-incrimination, and being content to resolve that issue.

335. One of the most familiar of judicial avoidance strategies is acting as though there really is no gatekeeping issue, but only issues of credibility or weight which are to be decided by the jury.

336. In most instances, the expert's assertion of valid expertise was sufficient to gain admissibility. Moreover, the *Frye* test formalizes faith in experts. And though the *Frye* test has the wisdom of asking judges to assess the consensus among the *community* of experts, in practice that is rarely done, and the court ends up relying on the "say-so" of the expert(s) before the court.

337. See Risinger et al., *supra* note 99, at 782.

338. That astrology would pass at least the narrow *Frye* test, and yet would never be admitted as expert evidence, is perhaps the exception that proves the rule. That is to say, the courts reveal themselves to have a finger on the pulse of the general or elite publics in order to know that they would invite ridicule by admitting expert astrological testimony, merely because those techniques were "generally accepted" among astrologers. They would not follow the *Frye* rule to the logical conclusion to admit astrology because to do so would not

Of course, scientific truth and what the public has faith in often will be one and the same thing. But where they differ, courts face the dilemma of either admitting untruths that are nevertheless widely subscribed to (thereby impeding the search for truth but enhancing the public's confidence in the courts) or excluding them (thereby advancing the search for truth but reducing the public's confidence in the courts). From this perspective, *Daubert* is a bold and risky departure from tradition. In its theory, at least, *Daubert's* approach places so high a value on truth-seeking that it is willing to risk the episodic (and perhaps cumulative) loss of public confidence.<sup>339</sup> On the other hand, by dedicating itself to filtering out any and all empirical claims that cannot be demonstrated to be valid, perhaps public confidence in the trial process will be enhanced in the long run.

Finally, in a world of continually advancing scientific knowledge and technological innovations, the courts are in essence presented with the choice of keeping pace or being dismissed as obsolete Luddites. Every time a new, asserted, and at least plausible scientific expertise presented itself, especially in the service of convicting criminal defendants, courts cast their lot with modernity. And we have seen courts giving themselves large pats on the back for doing exactly this, while displaying no more scientific acumen than the most superstitious shaman.

## B. The Law Can Affect the Growth of Science

The choice of admissibility rules and associated procedures may also have the purpose of influencing the nature or direction of the development of science. The courts can create incentives to do more and better research, or disincentives for doing so.

Fundamental new discoveries risk raising judicial doubts about all that had gone before, and what the future may reveal about the present. No advances means raising fewer doubts. We have seen examples of forensic identification sciences—fingerprints, handwriting, tool marks—that have been largely frozen in time, with little if any fundamental progress since their foundational appearances in court.<sup>340</sup> This

---

advance this arguable goal of the *Frye* test, to garner public confidence.

339. For example, if the public believes in the scientific and professional claims of handwriting experts, but the courts exclude it, the public will ask why the courts bury their heads in the sand, and refuse to accept the benefits that asserted expertise has to offer. In this light, *Starzecpyzel* is an ingenious opinion: It serves the scientific truth-seeking values of *Daubert* by declaring that handwriting expertise is not based on science, while still soliciting public support by recognizing that the public believes it will be helpful.

340. By fundamental I mean the basic scientific understanding these fields have of the phenomena with which they deal. I do not mean advances in mere technique; for example, new techniques for making latent prints visible or for photographing writings.



freezing resulted from a combination of premature (or at least science-free) judicial acceptance, the lack of other institutions (such as industry or academia) where competition or critical evaluation might create incentives for improved knowledge as well as improved technique. It is interesting to wonder what effect the competition of the adversary process might have had, if the opponents of admission had had the resources to mount challenges with the help of equal and opposing experts, challenges which the proponent's experts would then have had to meet.

The response of the courts also can inflate a field's views of its own capabilities. Recall, in the instance of bite mark identification, the courts did more to convince the experts that their knowledge was up to the task of pinpoint identification than forensic odontologists did to convince the courts. Such encouragement can slow progress at the same time that it enlarges a field's belief in itself. Recall the observation that "[t]he wholesale acceptance, by the courts, of testimony on bite mark identification has transformed the profession. . . . Having received judicial approval of bite mark comparisons, there seems to be no more limit on the extent of forensic odontological conclusions."<sup>341</sup>

There is no reason to think that the courts could not have a more productive effect on the growth of knowledge.<sup>342</sup> The Supreme Court's decision in *Daubert* prompted forensic scientists to begin to re-examine the supportability of their fields' claims.<sup>343</sup> *Starzecpyzel*, in particular, led directly to invitations from handwriting experts to some of their critics to meet and discuss both what is deficient in the underlying support for the claims of handwriting identification expertise and what research might be undertaken to begin, at long last, to install a scientific foundation under their work.<sup>344</sup>

Explicit efforts by the law to promote improved quality of forensic science have been suggested in the form of regulation<sup>345</sup> or new organizational arrangements.<sup>346</sup> It may well be that admissibility decisions of

---

341. MOENSSENS AT AL., *supra* note 4, § 16.07, at 985.

342. For an example of such impact outside of forensic science, see Joseph Sanders, *The Bendectin Litigation: A Case Study in the Life Cycle of Mass Torts*, 43 HASTINGS L.J. 301 (1992) (the life cycle of mass toxic tort litigation includes the growth of research on the effects of the suspected toxin following early litigation asserting a causal connection between the chemical and the plaintiff's injuries).

343. The 1996 annual meeting of the American Academy of Forensic Sciences had three major sessions in several different areas of forensic identification devoted to the challenges posed by *Daubert*. See PROCEEDINGS OF THE AAFS, 1996, NASHVILLE, TN (see, e.g., sessions E1, E2, and E3).

344. See *id.*

345. See Randolph N. Jonakait, *Forensic Science: The Need for Regulation*, 4 HARV. J.L. & TECH. 109 (1991).

346. See Giannelli, *supra* note 104.

courts are instruments too blunt to guide the development of scientific fields. Moreover, their very bluntness apparently can produce too little or the wrong kind of research by a field, at least fields that are oriented almost exclusively toward the courts and have the benefit of no other, more constructive, institutional pressures.<sup>347</sup>

### C. Sometimes the Law has no Effect on the Law

The experience in this area of law reveals a number of important cases where the law failed to have the expected impact on itself. This is worth documenting in its own right. But we also can ask whether the evasion illuminates some purpose behind the rule which is (or is not) present.

#### (1) *Frye can be Evaded*

On the same day that he published *Frye*, Judge Van Orsdel ignored his own test in another scientific evidence case of first impression. It should be no wonder then, that other judges found ways to evade the test, even though the law of their jurisdiction required them to use it, such as in *Marx*, the cornerstone case of forensic odontology. The voice spectrography cases provide an excellent illustration of the manipulability of the *Frye* test: To set the stage so that the proffered science is admitted, define the "particular field in which it belongs" narrowly, consisting of those who practice and promote the asserted science or technology. To set the stage for exclusion, define the field broadly, including those allied fields with more rigorous standards than those of the technique's practitioners.<sup>348</sup>

#### (2) *Daubert can be Evaded*

*Daubert* might have been thought to exclude handwriting identification by forensic document examiners on a court's finding that no data supported the proponents' claims of expertise, along with data that raised severe doubts about examiners' assumptions and abilities. Such findings by a court would seem to be the opposite of establishing evidentiary reliability. But the *Starzecpyzel* court, as well as several

---

347. For example, limits exist as to what engineers or physicians or psychiatrists can say in court because they have academic and commercial constituencies to whom they must answer. Their academic and scientific peers demand accuracy and their commercial clients demand progress. And discrepancies between what is asserted in court and what has been learned due to the expectations of those other constituencies cannot grow too large. Nothing comparable exists for the forensic identification sciences.

348. This is a flaw in the *Frye* test well recognized in the legal commentary on the subject. See Paul Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, A Half Century Later*, 80 COLUM. L. REV. 1197 (1980).

others that followed in its wake, re-classified that offering as non-science, thereby freeing themselves to apply a different and far more porous test of admissibility. Ironically, as a result of *Starzecpyzel* and its progeny, it can be said that courts had once admitted handwriting identification expertise because they believed it was scientifically well-founded and well-tested, while now they admit it because they know it is scientifically unfounded and untested.

(3) *Fundamental Principles can be Evaded*

If nothing else, one might have thought that inherent in the act of making a ruling is the requirement of writing a judicial opinion in which the court explains its reasoning. Indeed, the legitimacy of judicial authority depends, at a minimum, upon giving reasons for the decisions.<sup>349</sup> Absent reasons, the law is rendered arbitrary and courts surrender their authority. Yet in the cases we have examined, judges are often at a loss to explain their reasons, perhaps because they have so little understanding of the scientific evidence on which they are ruling that they do not even pretend to explain their reasoning in their opinions.<sup>350</sup> An especially clear example is provided by the two foundational tool mark cases of the Supreme Court of Washington State. The second of these opinions reversed the holding of the first, affirmed six months earlier, in cases involving the same expertise and the same expert. With respect to the asserted science of tool mark identification, the court did not explain the change in its understanding that led to its reversal of position. Nor did it explain the defects it believed it saw in the asserted science that led, only months before, to rejecting tool mark identification expertise.<sup>351</sup>

These cases, in numerous different areas of forensic identification science and in diverse ways, suggest a tension between the goals of the judges and the goals of the admissibility rules. Or perhaps the rules are constructed to be flexible so as to allow more intuitive, more popular, expertise-by-expertise decisions to be made. What seems clear from these cases is that no rule is capable of preventing admission when a judge is inclined to admit; each rule permits the court ample room to squeeze testimony through the filter.

---

349. See STEVEN J. BURTON, AN INTRODUCTION TO LAW AND LEGAL REASONING 167-71 (1985); KARL N. LLEWELYN, THE COMMON LAW TRADITION: DECIDING APPEALS 26 (1960).

350. Perhaps because what they have experienced is intuition and emotions, or at least trust, but no mature reasoning.

351. See *supra* notes 193-212 and accompanying text.

#### D. Courts Mistake Controversy for Error and Consensus for Validity

In many of the cases we have reviewed, the courts were presented with only one-sided questions regarding the adequacy of a given kind of asserted scientific evidence. Prosecutors typically offered the novel forensic science and defendants typically offered no reply of substance. The courts in these cases often said they were impressed at the "uncontradicted" expert testimony.

The adversary process is founded on the assumption that the truth of contestable propositions is most likely to emerge from structured disagreement. This fundamental belief seems to be forgotten as soon as courts try to evaluate science. Courts seem to expect other fields to resolve their uncertainties and reach a conclusion without disputes and without conflicting information. More than that, courts seem to expect of science what no earthly discipline can deliver, namely, something approaching certainty.<sup>352</sup> And more than that, courts appear to interpret the absence of controversy as the principal indication that infallible truth has been achieved. In part because *Daubert* requires, or at least invites, a more adversarial inquiry into the substance of science, judges are about to discover a world of science that is far less infallible than their past approaches allowed them to regard it as being.<sup>353</sup>

Rather than viewing vigorous debate within a scientific field as a healthy sign that uncertainties are being addressed, the courts seem to mistake debate for the presence of error. By contrast, would Anglo-American judges have more confidence in the findings of fact that emerged from a trial in which only one side was presented and the fact finder reached an unconflicted decision based on a one-sided hearing? Or more confidence in the findings of a trial in which decision-making was informed by contrasting evidence and vigorous debate?

Lack of controversy in science may well signal that a proposition of importance to the particular field has been thoroughly tested and settled. But it may also signal the absence of vigorous inquiry, an impoverished research tradition, lack of resources, or stagnation. Moreover, in criminal courts, the absence of countervailing experts probably reflects simply the lack of the resources necessary to support an adversarial contest. On the civil side, corporate titans rarely are at a loss to

---

352. See Imwinkelried, *supra* note 153. But rather than holding scientific evidence to impossibly high standards and excluding it, courts usually admit it, and in so doing tend to attribute to it higher levels of accomplishment than it has achieved. Not: We admit the evidence because it is excellent. But rather: Since we are going to admit it it must be excellent.

353. See *id.* Also consider the incident of the judge who had berated a physician for testifying in court about the field's belief about signs of injury, which the physician now was explaining to this meeting of judges was no longer considered valid by her field. See note 111, *supra*. On the general subject of the law's ability to deconstruct science, see Sheila Jasanoff, *What Judges Should Know About the Sociology of Science*, 32 JURIMETRICS J. 345 (1992).

contest any and every issue, scientific or otherwise.

Adversarial disagreement can be the means for challenging junk science and alerting a court to it. Silence, lack of disagreement, might not signal truth or even informed consensus. The structural absence of controversy is no assurance that truth has been found.

### E. Judges Do Not Think Like Scientists

Just as legal training teaches one the intellectual skills to analyze legal problems, scientific training teaches one how to analyze empirical questions and proposed answers. This places judges in a weak position to know what questions need to be asked in order to test an empirical claim or how to evaluate the data offered in answer.<sup>354</sup> *Daubert* requires judges to think like scientists. Some of the cases we have reviewed offer a preview of the struggles that judges must face as they undertake this challenge for which they are untrained.

The foundational opinions evaluating fingerprints, tool marks, and the earlier handwriting cases show no serious, substantive discussion of the phenomenon at issue in the claimed expertise.<sup>355</sup> Sometimes those courts were silent on the assertedly scientific substance. Sometimes they were content merely to cite to publications in which someone else had endorsed the discipline, even if those endorsements contained little or no empirical evidence or analysis. Frequently, courts were satisfied that the proffered expert witness testified that the asserted science was valid or that the witness's views were generally accepted in the field, or both. But rarely, if ever, did they rely on substantive evaluations of theory and data.

These judges were, of course, doing nothing other than behaving like the lawyers they were trained to be. There is no reason to expect courts to display any aptitude for evaluating the quality of scientific evidence. Their personnel—judges and lawyers—have been prepared for a field on the other side of the canyon that separates science from the humanities. By training, they barely know where to begin thinking

---

354. I once asked a class of judges I was teaching to imagine that they were surgeons who wanted to resolve the field's disagreement about which of two alternative surgical procedures was better. How would they approach solving that problem? They sat in pained silence. One judge finally suggested that they should ask an expert. I reminded them that in this exercise they were the experts. How did they think experts find answers? Does God whisper the answers into their ears?

355. In recent handwriting identification cases, the courts recognized that forensic document examiners failed the *Daubert* test, and so they excused them from having to pass it, imagining that they somehow acquired a practical skill from spending years looking at writings. These courts apparently did not realize that the hypothesis of practical skill is itself an empirical claim, which can be tested, and can only be tested, with relevant empirical data, not with analogies.

about the problem before them. And the tools courts are provided with—the available rules and procedures—are not up to the task.<sup>356</sup>

But let us concede that the structure of our legal system requires generalist judges guided by generalist lawyers to make these decisions about the admissibility of scientific evidence.<sup>357</sup> That has to be the starting point from which the law tries to construct a system that can help courts do this job better.

**F. Lawyers and Judges can Manage to Overlook Even the Best and Most Accessible Knowledge in an Unfamiliar Field**

We have only one major “test” of this proposition, but its results can only frighten anyone who realizes what it suggests—a failure of the adversary process to do precisely what is claimed to be one of its principal virtues.

The National Academy of Sciences was asked by the FBI to evaluate voice spectrography used for the purpose of identifying suspects, and the Academy assembled a diverse and first-rate panel of experts to examine the scientific evidence on the question. The Academy published a detailed report of their conclusions, which the FBI promptly adverted to. Lawyers in trials around the country failed to find and bring the report to the attention of judges, judges failed to find the report, and several courts which clearly knew of the report failed to learn from it.<sup>358</sup>

In short, the adversary process failed to motivate lawyers to find and offer the most important evidence on the subject at issue. And, although judges making a ruling on admissibility as a matter of law<sup>359</sup> have a duty to find the necessary evidence themselves, the courts failed to do so in almost every case following the publication of this report.<sup>360</sup>

An approximate equivalent of this would be if lawyers and judges all managed to overlook, or to read and disregard, a controlling Supreme Court opinion on a point central to a case at hand. It may happen occasionally. But ten times out of eleven?

That scientific literature is less accessible to lawyers and judges than cases or statutes may be the explanation, but cannot be a justification.

---

356. It would never occur to anyone to ask scientists, without law training, to write judicial opinions. And if they were asked, we would not be surprised to find them missing the obvious, making major mistakes, and generally mucking up the job.

357. Unless legislatures decide categories of it for them, requiring the wholesale admission or exclusion of various types of asserted expertise. This, of course, offers no guarantee of scientific validity, only democratic acceptability.

358. They probably failed to read it.

359. See Walker & Monahan, *supra* note 51, at 497.

360. See *id.* See also FED. R. EVID. 201 advisory committee's note.

### G. Bad Science Makes Bad Law

Some fields are bubbling cauldrons of inquiry. No assumption passes unexamined. No claim goes untested. No test goes unrepeated. Members of the field are bred on skepticism. Nothing becomes a shrine in the temple of that field's knowledge unless and until it has passed the most demanding scrutiny. Other fields are far less rigorous.<sup>361</sup>

This contrast suggests one of the more subtle defects in the *Frye* test: General acceptance will be attained in less scientifically vigorous fields more readily than in more demanding fields. As a result, the courts will welcome the more dubious from weaker fields of science sooner than the more reliable from stronger fields. By contrast, under a *Daubert* type analysis, fields that engage in less vigorous and less rigorous research will offer the courts less research to examine, and the research that is offered will be of lower quality. In short, "scientific" fields without a strong scientific research tradition will make it harder for the courts to evaluate their offerings.<sup>362</sup> Courts forced to make decisions about the admissibility of science based on a small body of poorly designed studies (or no studies at all) are going to have a much harder time making good decisions.

### H. The Two Most Prominent Rules of Admissibility are Simply Markers of the Options

The principal admissibility rules—*Frye* and *Daubert*—are not so much alternative solutions to the challenge of legal gatekeeping of purported scientific evidence as they are the markers of the two extreme possibilities. *Frye* requires massive deference to extra-judicial authorities and *Daubert* requires massive do-it-yourself. *Frye* does not work because its measure of validity is the judgment of "the field," and the field may consist of nonsense. For example, the *Frye* doctrine cannot exclude astrology. *Frye* could work, but only if the fields whose wares were offered to courts were limited to those that were valid and vigorous scientific enterprises. Then the courts could have some confi-

---

361. For example, consider the following statement concerning a purported finding that artificial chemicals in the environment mimic human hormones and are causing harmful health effects: "But several leading scientists view such propositions as premature at best. [The asserted conclusion is] fueled more by hyperbole than facts. . . . [M]any of the claims . . . when examined, turn out to be a house of cards." "These scientists say they are not arbitrarily dismissing [the claims, but] there is a difference between a hypothesis and evidence." Gina Kolata, *Chemicals that Mimic Hormones Spark Alarm and Debate*, N.Y. TIMES, Mar. 19, 1996, at B5. Instant internal challenges of this sort are common in many fields, but are rare in the history of the forensic identification sciences.

362. See, e.g., *Starzecpyzel*, *supra* note 44.

dence that a consensus reached in those fields derived from the benefits we expect the methods and culture of science to produce. But that initial determination is essentially the same unsolved problem moved to a different stage in the decision process.

*Daubert* avoids the problems of *Frye* by making courts evaluate science in much the way that science evaluates science. But the evaluators are still lay judges. And judges are lawyers. And lawyers were smart kids who disliked math and science. So they went to law school. Can we really expect judges to serve as Solomons of manifold fields that they may never have heard of until the day an issue, perhaps answered by one of those fields, arrives in their court?<sup>363</sup>

But perhaps the purpose of the rules is simply to hold up a target to the courts; call one the *Frye* target and the other the *Daubert* target. The *Frye* ideal says: do whatever the experts tell you to do. The *Daubert* ideal says: figure out the science yourself. No court really is expected to or desired to be as deferential as *Frye* would seem to require or as all-knowing as *Daubert* would seem to require. Courts are expected, instead, to make a reasonable, even if intuitive and unarticulable decision, that is a nearer or more distant approximation of the target adopted by its jurisdiction. But courts of both jurisdictional types are, realistically, expected to fall somewhere between the two targets, and never actually to hit either extreme.

The search for workable and effective legal tests of scientific evidence will have to continue. Perhaps effective solutions are more likely to be found in systems of procedure than in rules of admissibility,<sup>364</sup> perhaps even to the point of developing new institutions or forums within or alongside of the existing courts.<sup>365</sup> Whether the courts' management of the science that comes knocking on their door will ever improve is a question that only the future can answer. In the meantime, society can ask of its judges the same question that Earl Weaver asked of the umpires in his field: "Are you gonna get any better, or is this it?"<sup>366</sup>

---

363. Justice Blackmun says yes. Chief Justice Rehnquist says no. I point to the cases cited throughout this article as one source of intelligence about how judges perform at these tasks. If these cases provide any foresight on the question, many judges will do what they can to make the task easy on themselves.

364. See Samuel R. Gross, *Expert Evidence*, 1991 WISC. L. REV. 1113, for suggested solutions.

365. See, e.g., Arthur Kantrowitz, *The Science Court Experiment*, 7 JURIMETRICS J. 332 (1977).

366. Quoted by George Will, *LBJ Administration Redux*, DALLAS MORNING NEWS, May 17, 1993, at 15A.



## Appendix: Scientific Voice Identification Cases

### Holdings, Legal Tests, and Citations to NAS Report

Jurisdiction	Court	Case	Cite	Date	Legal Test*	Held	NAS Report
Military	APP	Wright	17 CMA 183	1967	Reliability	IN	
CA	APP	King	72 Cal. Rptr. 478	1968	Frye-broad	OUT	
NJ	TR	Cary	239 A.2d 680	1970	Frye-broad	OUT	
MN	SC	Trimble	192 N.W.2d 432	1971	none	IN	
FL	APP	Worley	263 So.2d 613	1972	Reliability	IN	
FL	APP	Alea	265 So.2d 96	1972	none + [Reliability]	IN	
CA	APP	Hodo	106 Cal. Rptr. 547	1973	Frye-narrow	IN	
US-DC Cir.	APP	Addison	498 F.2d 741	1974	Frye-broad	OUT	
CA	APP	Law	114 Cal. Rptr. 708	1974	Frye-broad	OUT	
US-EDPA	TR	Sample	378 F. Supp. 44	1974	McC	IN	
MA	SC	Lykus	327 N.E.2d 671	1975	Frye-narrow	IN	
US-4 <sup>th</sup> Cir.	APP	Baller	519 F.2d 463	1975	McC	IN	
US-6 <sup>th</sup> Cir.	APP	Franks	511 F.2d 25	1975	McC + Reliability	IN	
US-6 <sup>th</sup> Cir.	APP	Jenkins	525 F.2d 819	1975	McC + [Reliability]	IN	
OH	APP	Olderman	44 Ohio App.2d 130	1975	Reliability	IN	
CA	SC	Kelly	130 Cal. Rptr. 144	1976	Frye-broad	OUT	
US-DC Cir.	APP	McDaniel	538 F.2d 408	1976	Frye-broad	OUT	
NY	TR	Rogers	385 N.Y.S.2d 228	1976	McC + Rel + [Frye]	IN	
PA	SC	Topa	369 A.2d 1277	1977	Frye-broad	OUT	
MI	SC	Tobey	257 N.W.2d 537	1977	Frye-broad	OUT	
US-SDNY	TR	Williams	443 F. Supp. 269	1977	Reliab + [Frye-broad]	IN	
MD	SC	Reed	391 A.2d 364	1978	Frye-broad	OUT	
US-2 <sup>nd</sup> Cir.	APP	Williams	583 F.2d 1194	1978	McC	IN	
NJ	TR	D'Arc	385 A.2d 278	1978	Reliab or Frye-broad	OUT	
DC	APP	Brown	384 A.2d 647	1978	[Reliability + Frye]	neither	
NY	TR	Collins	405 N.Y.S.2d 365	1978	Reliab + Frye-broad	OUT	
ME	SC	Williams	388 A.2d 500	1978	Reliability-Relevancy	IN	
NY	TR	Bein	453 N.Y.S.2d 343	1982	Reliab + Frye-narrow	IN	no
IN	SC	Cornett	450 N.E.2d 498	1983	Frye-broad	OUT	no

Jurisdiction	Court	Case	Cite	Date	Legal Test*	Held	NAS Report
OH	SC	Williams	446 N.E.2d 444	1983	Reliability	IN	no
AZ	SC	Gortarez	686 P.2d 1224	1984	Frye-broad	OUT	yes
RI	SC	Wheeler	496 A.2d 1382	1985	McC	IN	no
LA	APP	Free	493 So.2d 781	1986	Relevancy balance	OUT	slightly
NJ	SC	Windmere	522 A.2d 405	1987	Frye-broad	OUT	no
CO	SC	Drake	748 P.2d 1237	1988	Frye-broad	OUT	no
US-7 <sup>th</sup> Cir.	APP	Smith	869 F.2d 348	1989	Reliability + [Frye]	IN	slightly
US-DHI	TR	Maivia	728 F. Supp. 1471	1990	Reliab + Frye-narrow	IN	slightly
US-6 <sup>th</sup> Cir.	APP	Leon	966 F.2d 1455 (Table)	1992	McC	IN	no

The legal tests are abbreviated as follows: Frye with the relevant fields defined broadly (Frye-broad), or narrowly (Frye-narrow), reliability (reliab) or relevancy (relev), or McCormick weighting (McC).

The court levels are: court of last resort (sc), court of appeals (ca), trial (tr). Brackets indicate a test a court stated it was applying but where there is no indication in the opinion that the court actually applied that test.

